

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: October 7, 2003, 07:04:18 : Search time 96 Seconds  
(without alignments)  
908.560 Million cell updates/sec

Title: US-09-856-679-2  
Perfect score: 1759  
Sequence: 1 MAALQEDGWTGQVLKVN.....OLKVIDNQRELSRLRELEP 338

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mhc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_rodent: \*  
12: sp\_virus: \*  
13: sp\_vertebrate: \*  
14: sp\_unclassified: \*  
15: sp\_virus: \*  
16: sp\_bacteriophage: \*  
17: sp\_archaeophage: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	1759	100.0	881	4	Q8WVNO	Q8WVNO homo sapien
2	1755	99.8	881	4	Q95398	Q95398 homo sapien
3	1748	99.4	881	4	Q95634	Q95634 homo sapien
4	1664	94.6	884	11	Q921C8	Q921C8 rattus norv
5	1655	94.1	461	11	Q8BZK9	Q8BZK9 mus musculu
6	1591	90.4	876	11	Q8VCC8	Q8VCC8 mus musculu
7	1432	80.8	287	11	Q8R1R1	Q8R1R1 mus musculu
8	863.5	49.1	1011	4	Q8WZAZ	Q8WZAZ homo sapien
9	863.5	49.1	1011	4	Q95636	Q95636 homo sapien
10	857.5	48.7	1011	4	Q8TAA4	Q8TAA4 homo sapien
11	855.5	48.6	696	11	Q8VIR9	Q8VIR9 mus musculu
12	855.5	48.6	699	11	Q9CWS2	Q9CWS2 mus musculu
13	855.5	48.6	867	4	Q8IXK6	Q8IXK6 homo sapien
14	855.5	48.6	993	11	Q921P0	Q921P0 mus musculu
15	855.5	48.6	1011	11	Q9E0Z6	Q9E0Z6 mus musculu
16	832.5	47.3	436	11	Q921C7	Q921C7 rattus norv

17	696	39.6	861	5	Q9V9A3	Q9V9A3 drosophila
18	655	37.2	580	4	Q92565	Q92565 homo sapien
19	644.5	36.6	612	11	Q8B1J9	Q8B1J9 mus musculu
20	644.5	36.6	814	11	Q8C0Q9	Q8C0Q9 mus musculu
21	637.5	36.2	456	4	Q9UHV5	Q9UHV5 homo sapien
22	637.5	36.2	814	11	Q8C0R5	Q8C0R5 mus musculu
23	495.5	28.2	444	4	Q8IXU5	Q8IXU5 homo sapien
24	459	26.1	1311	5	Q95NLS	Q95NLS caenorhabd
25	459	26.1	1347	5	Q95WR8	Q95WR8 caenorhabd
26	459	26.1	1470	5	Q21218	Q21218 caenorhabd
27	409	23.3	1499	4	Q914G8	Q914G8 homo sapien
28	406	23.1	1138	11	Q8CHG7	Q8CHG7 mus musculu
29	404.5	23.0	1573	5	Q95V18	Q95V18 drosophila
30	402.5	22.9	1573	5	Q9VWF3	Q9VWF3 drosophila
31	390.5	22.2	1204	4	Q9UHV4	Q9UHV4 homo sapien
32	390.5	22.2	1391	4	Q8TEU6	Q8TEU6 homo sapien
33	390.5	22.2	1601	4	Q8TEU7	Q8TEU7 homo sapien
34	390.5	22.2	1601	4	Q8N121	Q8N121 homo sapien
35	387.5	22.0	834	11	Q8R3E5	Q8R3E5 mus musculu
36	386.5	22.0	1509	4	Q96PC1	Q96PC1 homo sapien
37	380.5	21.6	1113	4	Q8REA3	Q8REA3 homo sapien
38	306.5	17.4	929	5	Q8SSQ3	Q8SSQ3 dicystostell
39	304.5	17.3	1721	5	Q8SSQ0	Q8SSQ0 dicystostell
40	301.5	17.1	824	5	Q8IS15	Q8IS15 dicystostell
41	301	17.1	812	5	Q8IS14	Q8IS14 dicystostell
42	283.5	16.1	642	11	Q8C5V7	Q8C5V7 mus musculu
43	283.5	16.1	1086	11	Q91Z22	Q91Z22 mus musculu
44	283	16.1	1095	4	Q8IV73	Q8IV73 homo sapien
45	279	15.9	1013	11	Q9QYV3	Q9QYV3 rattus norv

#### ALIGNMENTS

RESULT 1  
Q8WVNO PRELIMINARY; PRT; 881 AA.  
ID Q8WVNO  
AC Q8WVNO: 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Rapi guanine-nucleotide-exchange factor directly activated by  
DE CAMP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxId=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Strausberg R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
EMBL: BC017728; AAH17728.1; -  
DR InterPro: IPR002373; CAMP\_kin.  
DR InterPro: IPR000595; CNMP\_binding.  
DR InterPro: IPR000591; DEP.  
DR InterPro: IPR000651; RASGEFN.  
DR InterPro: IPR001895; RASGEFN\_CDC25.  
DR Pfam: PF00027; CNMP\_binding; 1.  
DR Pfam: PF00610; DEP; 1.  
DR Pfam: PF00617; RASGEFN; 1.  
DR Pfam: PF00618; RASGEFN; 1.  
DR PRINTS: PRO0103; CAMPKINASE.  
DR SMART: SM00100; CNMP; 1.  
DR SMART: SM0049; DEP; 1.  
DR SMART: SM00147; RASGEFN; 1.  
DR SMART: SM00229; RASGEFN; 1.  
DR PROSITE: PS00042; CNMP\_BINDING\_3; 1.  
DR PROSITE: PS00186; DEP; 1.  
SQ SEQUENCE 881 AA; 99352 MW; D1045DZADPFB29A CRC64;

Query Match 100.0%; Score 1759; DB 4; Length 881;  
Best Local Similarity 100.0%; Pred. No. 5e-152;

Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAALAEDGWTGKGVLYKVNAGDAIGLQDPARGVATSLGNERLFVYNPOEVHLLIHP 60  
|||||  
DB 544 MAALAEDGWTGKGVLYKVNAGDAIGLQDPARGVATSLGNERLFVYNPOEVHLLIHP 603  
|||||  
QY 61 DQGPVTGASAGLDLVSAKDLAAGLTDHDSLFNSIHQVELIHVVLGPQHLRDVTANLE 120  
|||||  
DB 604 DQGPVTGASAGLDLVSAKDLAAGLTDHDSLFNSIHQVELIHVVLGPQHLRDVTANLE 663  
|||||  
QY 121 RFMRPFNELOYWATELCLCPVPGPRAQLLRKFKTLAAHLKEOKNLNSFFAVMGLSNSA 180  
RFMRPFNELOYWATELCLCPVPGPRAQLLRKFKTLAAHLKEOKNLNSFFAVMGLSNSA 723  
DB 664 RFMRPFNELOYWATELCLCPVPGPRAQLLRKFKTLAAHLKEOKNLNSFFAVMGLSNSA 723  
|||||  
QY 181 ISRLAHMERLPHVVRKLYSALERLDPSSMNHRYRLALAKSPVPIPFMPLLDKMTFI 240  
|||||  
DB 724 ISRLAHMERLPHVVRKLYSALERLDPSSMNHRYRLALAKSPVPIPFMPLLDKMTFI 783  
|||||  
QY 241 HEGNHTLVENLINFEKKRMARAARMLHCRSHNPVPLSPRSVSHLHEDSOVARISTC 300  
|||||  
DB 784 HEGNHTLVENLINFEKKRMARAARMLHCRSHNPVPLSPRSVSHLHEDSOVARISTC 843  
|||||  
QY 301 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 338  
|||  
DB 844 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 881  
|||

RESULT 2  
095398 PRELIMINARY; PRT: 881 AA.

AC 095398:  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Rapl guanine-nucleotide exchange factor.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Muscle;  
RX MEDLINE=99066645; PubMed=9853756;  
RA de Rooij J., Zwartkruis F.J., Verheijen M.H., Cool R.H., Nijman S.M.,  
Wittinghofer A., Bos J.L.;  
RT "Epac is a Rapl guanine-nucleotide-exchange factor directly activated  
RT by cyclic AMP.";  
RL Nature 396:474-477(1998).  
DR EMBL: AF103905; AAC83381.1; -  
DR InterPro: IPR002373; CAMP\_kin.  
DR InterPro: IPR000595; CNMP\_binding.  
DR InterPro: IPR000651; DEP.  
DR InterPro: IPR001895; RasGEFN.  
DR InterPro: IPR001895; RasGEFN\_CDC25.  
DR Pfam: PF00027; CNMP\_binding; 1.  
DR Pfam: PF00610; DEP; 1.  
DR Pfam: PF00617; RasGEFN; 1.  
DR Pfam: PF00618; RasGEFN; 1.  
DR PRINTS: PR00103; CAMPKINASE.  
DR SMART: SM00100; CNMP; 1.  
DR SMART: SM00049; DEP; 1.  
DR SMART: SM00147; RasGEFN; 1.  
DR SMART: SM00229; RasGEFN; 1.  
DR PROSITE: PS50042; CNMP\_BINDING\_3; 1.  
DR PROSITE: PS50186; DEP; 1.  
SQ SEQUENCE 881 AA; 99312 MW; 76B82C54E0316D45 CRC64;

Query Match 99.8%; Score 1755; DB 4; Length 881;  
Best Local Similarity 99.7%; Pred. No. 1.2e-151;  
Matches 337; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAALAEDGWTGKGVLYKVNAGDAIGLQDPARGVATSLGNERLFVYNPOEVHLLIHP 60  
|||||

DB 544 MAALAEDGWTGKGVLYKVNAGDAIGLQDPARGVATSLGNERLFVYNPOEVHLLIHP 603  
|||||  
QY 61 DQGPVTGASAGLDLVSAKDLAAGLTDHDSLFNSIHQVELIHVVLGPQHLRDVTANLE 120  
|||||  
DB 604 DQGPVTGASAGLDLVSAKDLAAGLTDHDSLFNSIHQVELIHVVLGPQHLRDVTANLE 663  
|||||  
QY 121 RFMRPFNELOYWATELCLCPVPGPRAQLLRKFKTLAAHLKEOKNLNSFFAVMGLSNSA 180  
RFMRPFNELOYWATELCLCPVPGPRAQLLRKFKTLAAHLKEOKNLNSFFAVMGLSNSA 723  
DB 664 RFMRPFNELOYWATELCLCPVPGPRAQLLRKFKTLAAHLKEOKNLNSFFAVMGLSNSA 723  
|||||  
QY 181 ISRLAHMERLPHVVRKLYSALERLDPSSMNHRYRLALAKSPVPIPFMPLLDKMTFI 240  
|||||  
DB 724 ISRLAHMERLPHVVRKLYSALERLDPSSMNHRYRLALAKSPVPIPFMPLLDKMTFI 783  
|||||  
QY 241 HEGNHTLVENLINFEKKRMARAARMLHCRSHNPVPLSPRSVSHLHEDSOVARISTC 300  
|||||  
DB 784 HEGNHTLVENLINFEKKRMARAARMLHCRSHNPVPLSPRSVSHLHEDSOVARISTC 843  
|||||  
QY 301 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 338  
|||  
DB 844 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 881  
|||

RESULT 3  
095634 PRELIMINARY; PRT: 881 AA.

AC 095634:  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE CAMP-regulated guanine nucleotide exchange factor I.  
GN CAMP-GEFI.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99074384; PubMed=9856955;  
RA Kawasaki H., Springett G.M., Mochizuki N., Toki S., Nakaya M.,  
Matsuda M., Housman D.E., Graybiel A.M.;  
RT "A family of CAMP-binding proteins that directly activate rapl.";  
RL Science 282:2275-2279(1998).  
DR EMBL: U78168; AAD12740.1; -  
DR InterPro: IPR000595; CNMP\_binding.  
DR InterPro: IPR000651; DEP.  
DR InterPro: IPR001895; RasGEFN.  
DR InterPro: IPR001895; RasGEFN\_CDC25.  
DR Pfam: PF00027; CNMP\_binding; 1.  
DR Pfam: PF00610; DEP; 1.  
DR Pfam: PF00617; RasGEFN; 1.  
DR Pfam: PF00618; RasGEFN; 1.  
DR SMART: SM00100; CNMP; 1.  
DR SMART: SM00049; DEP; 1.  
DR SMART: SM00147; RasGEFN; 1.  
DR SMART: SM00229; RasGEFN; 1.  
DR PROSITE: PS50042; CNMP\_BINDING\_3; 1.  
DR PROSITE: PS50186; DEP; 1.  
SQ SEQUENCE 881 AA; 99376 MW; 971CA7A6EEF7B1E CRC64;

Query Match 99.4%; Score 1748; DB 4; Length 881;  
Best Local Similarity 99.1%; Pred. No. 5.1e-151;  
Matches 335; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAALAEDGWTGKGVLYKVNAGDAIGLQDPARGVATSLGNERLFVYNPOEVHLLIHP 60  
|||||  
DB 544 MAALAEDGWTGKGVLYKVNAGDAIGLQDPARGVATSLGNERLFVYNPOEVHLLIHP 603  
|||||  
QY 61 DQGPVTGASAGLDLVSAKDLAAGLTDHDSLFNSIHQVELIHVVLGPQHLRDVTANLE 120  
|||||  
DB 604 DQGPVTGASAGLDLVSAKDLAAGLTDHDSLFNSIHQVELIHVVLGPQHLRDVTANLE 663  
|||||  
QY 121 RFMRPFNELOYWATELCLCPVPGPRAQLLRKFKTLAAHLKEOKNLNSFFAVMGLSNSA 180  
RFMRPFNELOYWATELCLCPVPGPRAQLLRKFKTLAAHLKEOKNLNSFFAVMGLSNSA 723  
DB 664 RFMRPFNELOYWATELCLCPVPGPRAQLLRKFKTLAAHLKEOKNLNSFFAVMGLSNSA 723  
|||||  
QY 181 ISRLAHMERLPHVVRKLYSALERLDPSSMNHRYRLALAKSPVPIPFMPLLDKMTFI 240  
|||||  
DB 724 ISRLAHMERLPHVVRKLYSALERLDPSSMNHRYRLALAKSPVPIPFMPLLDKMTFI 783  
|||||  
QY 241 HEGNHTLVENLINFEKKRMARAARMLHCRSHNPVPLSPRSVSHLHEDSOVARISTC 300  
|||||  
DB 784 HEGNHTLVENLINFEKKRMARAARMLHCRSHNPVPLSPRSVSHLHEDSOVARISTC 843  
|||||  
QY 301 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 338  
|||  
DB 844 SEQSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 881  
|||

ID	Accession	Species	Gene	Length	Score	DB	Length	Score	DB
Db	664	RFMRFRNFELQYVWATELCLCPVCPGPRQALLKFKIKLAHLKEQKVNSEFFAVMFLSNP		723					
Qy	181	ISRLAHMERLPHKVRKLYSALERLLDPSWNHRYRLAKLSPVIDPMPLLKDMTFI		240					
Db	724	ISRLAHMERLPHKVRKLYSALERLLDPSWNHRYRLAKLSPVIDPMPLLKDMTFI		783					
Qy	241	HEGHHHTVENLINEKMMAMRAARMLHCHSHNPVPLSPRSRSHLHDSQVARI		300					
Db	784	HEGHHHTVENLINEKMMAMRAARMLHCHSHNPVPLSPRSRSHLHDSQVARI		843					
Qy	301	SEQSLSTRSPASTWAVYQQLAVINQRELSTRLELP		338					
Db	844	SEQSLSTRSPASTWAVYQQLAVINQRELSTRLELP		881					
RESULT 4									
ID	0921C8	PRELIMINARY:	PRF:	884 AA.					
AC	0921C8:								
DT	01-MAY-1999 (TREMBLrel. 10, Created)								
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)								
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)								
DE	CAMP-regulated guanine nucleotide exchange factor I.								
GN	CAMP-GEFI.								
OC	Rattus norvegicus (Rat).								
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.								
NCBI	TextID=10116;								
RP	11								
RA	SEQUENCE FROM N.A.								
RX	MEDLINE=99074384; PubMed=9856955;								
RA	Kawasaki H., Springle G.M., Mochizuki N., Toki S., Nakaya M.,								
RA	Matsuda M., Housman D.E., Graybiel A.M.;								
RT	"A family of CAMP-binding proteins that directly activate rap1."								
RL	Science 282:2275-2279(1998).								
DR	EMBL: U78167; AAD12739.1; "								
DR	InterPro: IPR002373; CAMP_kin								
DR	InterPro: IPR000593; cAMP_binding.								
DR	InterPro: IPR000591; DEP.								
DR	InterPro: IPR001895; RasGEF.								
DR	InterPro: IPR001895; RasGEF_CDC25.								
DR	Pfam: PF00027; cAMP_binding.1.								
DR	Pfam: PF00610; DEP.1.								
DR	Pfam: PF00617; RasGEF.1.								
DR	Pfam: PF00618; RasGEF.1.								
DR	PRINTS: PR00103; CAMPKINASE.								
DR	SMART: SM00100; cAMP_KINASE.								
DR	SMART: SM00049; DEP.1.								
DR	SMART: SM00147; RasGEF.1.								
DR	SMART: SM00229; RasGEF.1.								
DR	PROSITE: PS50042; cAMP_BINDING_3.1.								
DR	PROSITE: PS50186; DEP.1.								
SO	SEQUENCE 884 AA; 100256 MW; B082AAE3125AC933 CRC64;								
Query Match 94.6%; Score 1664; DB 11; Length 884;									
Best Local Similarity 93.5%; Pred. No. 2.5e-143;									
Matches 316; Conservative 11; Mismatches 11; Indels 0; Gaps 0;									
Qy	1	MAALAEQEGMKKGOVLVYNVNSAGDAIGQPRARGVATSLGNERLFTVYVNPQEVHETLHP		60					
Db	547	MAALAEHDMTKKGOVLVYNVNSAGDVGLQPRARGVATSLGNERLFTVVDPEVHETLHP		606					
Qy	61	DQLGPTVSGAGDLVSAKDLAGOLTDHDSLFNSIHQVELIHVYLGPOHLRDVTANLE		120					
Db	607	EQLGPTLGSSBMDLVSAKDLAGOLTEHDMNLFNHIQVELIHVYLGPOHLRDVTANLE		666					
Qy	121	RFMRFRNFELQYVWATELCLCPVCPGPRQALLKFKIKLAHLKEQKVNSEFFAVMFLSNP		180					
Db	667	RFMRFRNFELQYVWATELCLCPVCPGPRQALLKFKIKLAHLKEQKVNSEFFAVMFLSNP		726					
Qy	181	ISRLAHMERLPHKVRKLYSALERLLDPSWNHRYRLAKLSPVIDPMPLLKDMTFI		240					

ID	Q8BZK9	PRELIMINARY:	PRT:	461 AA.
Db	727	ISRLAHNTERLPKHYVRKLYSALERLDPSWMHRYRRLATLTGSLPPVIFPMPLLLKDMFI	786	
Qy	241	HEGHNHTLVENLINEFKMMARMAARMLHCHSHNPVPLSPLSRSYSHLHDSQVARISTC	300	
Db	787	HEGHNHTLVENLINEFKMMARMAARVMLHCHSHSHSTAPLSRPSVSHIHEDSQSARISTC	846	
Qy	301	SEQSILSTRSPASTWAVYQOLKVIDNQRLSRLSRELP	338	
Db	847	SEQSILSTRSPASTWAVYQOLKVIDNQRLSRLSRELP	884	
RESULT 5				
ID	Q8BZK9	PRELIMINARY:	PRT:	461 AA.
AC	Q8BZK9			
DT	01-MAR-2003	(TREMBLrel. 23, Created)		
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)		
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)		
DE	CAMP-regulated guanine nucleotide exchange factor I homolog (Fragment).			
OC	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.			
OX	NCBI_Taxid=10090;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Diencephalon;			
RX	MEDLINE=22354683; PubMed=12466851;			
RA	THE FANTOM Consortium,			
RA	THE RIKEN Genome Exploration Research Group Phase I & II Team;			
RT	"Analysis of the mouse transcriptome based on functional annotation of			
RT	60,770 full-length cDNAs."			
RL	Nature 420:563-573(2002).			
DR	EMBL; AK034265; BAC28653.1; ..			
FT	NON_TER			
FT	1			
FT	1			
SEQUENCE	461 AA; 52434 MW; 638101EPID2E4249 CRC64;			
Query Match 94.1%; Score 1655; DB 11; Length 461;				
Best Local Similarity 93.5%; Pred. No. 6.8e-143;				
Matches 316; Conservative 10; Mismatches 12; Indels 0; Gaps 0;				
Qy	1	MAALAOEGMFKRGQVLVYVNSAGDITGLQPARGVATSLGILNERLFVYNPOEHNELPH	60	
Db	124	MAALAHNEHMTKGOVLVYVNSGDVVGQPPDARGVATSLGILNERLFVVDPOEHNELPH	183	
Qy	61	DOLGPTVVSAGCLDLVSAKDLAQGLTDHDWSLFNSIHQVELIHVYLGPHQLRDYTTANLE	120	
Db	184	EOLGPTLGSSEMDLDVSAKDLAQGLTDHDWMFNHIVQELIHVYLGPHQLRDYTTANLE	243	
Qy	121	RFMRFRNELQYVATLCLCPVPGFRAQLLRKFKILAAHLKQKULNSFFAVMFGLSNSA	180	
Db	244	RFMRFRNELQYVATLCLCPVPGSRAQLLRKFKILAAHLKQKULNSFFAVMFGLSNSA	303	
Qy	181	ISRLAHNTERLPKHYVRKLYSALERLDPSWMHRYRRLAKISPRVITPFPMLLLKDMFI	240	
Db	304	ISRLAHNTERLPKHYVRKLYSALERLDPSWMHRYRRLATLTGSLPPVIFPMPLLLKDMFI	363	
Qy	241	HEGHNHTLVENLINEFKMMARMAARVMLHCHSHNPVPLSPLSRSYSHLHDSQVARISTC	300	
Db	364	HEGHNHTLVENLINEFKMMARMAARVMLHCHSHSTAPLSRPSVSHIHEDSQSARISTC	423	
Qy	301	SEQSILSTRSPASTWAVYQOLKVIDNQRLSRLSRELP	338	
Db	424	SEQSILSTRSPASTWAVYQOLKVIDNQRLSRLSRELP	461	
RESULT 6				
ID	Q8VCC8	PRELIMINARY:	PRT:	876 AA.
AC	Q8VCC8			
DT	01-MAR-2002	(TREMBLrel. 20, Created)		
DT	01-MAR-2002	(TREMBLrel. 20, Last sequence update)		
DT	01-OCT-2002	(TREMBLrel. 22, Last annotation update)		

DE Similar to cAMP-regulated guanine nucleotide exchange factor I (cAMP-  
DE GEF1).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Strausberg R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC020533; AAH20533.1; -  
DR InterPro: IPR002373; CAMP\_kin  
DR InterPro: IPR000593; cAMP\_binding.  
DR InterPro: IPR000591; DEP.  
DR InterPro: IPR000651; RasGEFN.  
DR InterPro: IPR001895; RasGRF\_CCC25.  
DR Pfam: PF00027; cAMP\_binding; 1.  
DR Pfam: PF00610; DEP; 1.  
DR Pfam: PF00617; RasGEF; 1.  
DR Pfam: PF00618; RasGEFN; 1.  
DR PRINTS: PR00103; CAMPKINASE.  
DR SMART: SM00100; cAMP; 1.  
DR SMART: SM00049; DEP; 1.  
DR SMART: SM00147; RasGEF; 1.  
DR SMART: SM00229; RasGEFN; 1.  
DR PROSITE: PS50042; cAMP\_BINDING\_3; 1.  
DR PROSITE: PS50186; DEP; 1.  
SQ SEQUENCE: 876 AA; 99230 MW; 7DC12B83F18F9AF8 CRC64;

Query Match 90.4%; Score 1591; DB 11; Length 876;  
Best Local Similarity 90.5%; Pred. No. 1.2e-136;  
Matches 306; Conservative 12; Mismatches 12; Indels 8; Gaps 1;

QY 1 MAALAEQEGMTKGVLYVNSAGDAIGQPDARGVATSLGNERLFVYNQVEHLLPH 60  
DB 547 MAALAEHDMTKGVLYVNSAGDVYGLQPDARGVATSLGNERLFVYNQVEHLLPH 606  
QY 61 DQLGPTVGSAGDLVSAKDLAGQLTDHWSLFNSIHQVELIHVYLGPOHLDTVTANLE 120  
DB 607 EQLGPTLGSSEMLDLVSAKDLAGQLTDHDMNLFNRHQVQ-----EHLRDTVTANLE 658  
QY 121 RMRFRNELQYVAVTELCPCVPGRRAQLLRKFTLAHLKEQKLNSEFFAVMFLSNSA 180  
DB 659 RMRFRNELQYVAVTELCPCVPGRRAQLLRKFTLAHLKEQKLNSEFFAVMFLSNSA 718  
QY 181 ISRLAHTWERLPHKRYKLYSALERLLDPSNMHRYRYRLALAKLSPVIFPMPLLDKMTFI 240  
DB 719 ISRLAHTWERLPHKRYKLYSALERLLDPSNMHRYRYRLALAKLSPVIFPMPLLDKMTFI 778  
QY 241 HEGNHTLVENLINFKRMMAARAAMLHCRSHNFPVPLSRVSHLHEDSQVARISTC 300  
DB 779 HEGNHTLVENLINFKRMMAARAAMLHCRSHNFPVPLSRVSHLHEDSQVARISTC 838  
QY 301 SEQSLSTRSPASTWAVYQOLKVIYDQRELSTRSLRLEP 338  
DB 839 SEQSLSTRSPASTWAVYQOLKVIYDQRELSTRSLRLEP 876

## RESULT 7

Q8R1R1 PRELIMINARY; PRT; 287 AA.  
AC Q8R1R1;  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Hypothetical 33.2 kDa protein (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;  
RA Strausberg R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC020311; AAH20311.1; -  
DR InterPro: IPR001895; RasGRF\_CDC25.  
DR Pfam: PF00617; RasGEF; 1.  
DR SMART: SM00147; RasGEF; 1.  
KM Hypothetical protein.  
FT NON\_TER  
SQ SEQUENCE 287 AA; 33165 MW; 87D9AF9505928C4B CRC64;

Query Match 80.8%; Score 1422; DB 11; Length 287;  
Best Local Similarity 94.1%; Pred. No. 7.6e-122;  
Matches 270; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 52 EVHELIPHPDQLGPTVGSAGDLVSAKDLAGQLTDHWSLFNSIHQVELIHVYLGPOHL 111  
DB 1 EVHELIPHPDQLGPTVGSAGDLVSAKDLAGQLTDHDMNLFNRHQVQHL 60  
QY 112 RDVTANLERNRFRNELQYVAVTELCPCVPGRRAQLLRKFTLAHLKEQKLNSEFFA 171  
DB 61 RDVTANLERNRFRNELQYVAVTELCPCVPGRRAQLLRKFTLAHLKEQKLNSEFFA 120  
QY 172 VMFGLSNSAIRLAHTWERLPHKRYKLYSALERLLDPSNMHRYRYRLALAKLSPVIFPM 231  
DB 121 VMFGLSNSAIRLAHTWERLPHKRYKLYSALERLLDPSNMHRYRYRLALAKLSPVIFPM 180  
QY 232 LLLKDMFTIHEGNHTLVENLINFKRMMAARAAMLHCRSHNFPVPLSRVSHLHED 291  
DB 181 LLLKDMFTIHEGNHTLVENLINFKRMMAARAAMLHCRSHNFPVPLSRVSHLHED 240  
QY 292 SQVARISTCSQSSTRSPASTWAVYQOLKVIYDQRELSTRSLRLEP 338  
DB 241 SQGSRISTCSQSSTRSPASTWAVYQOLKVIYDQRELSTRSLRLEP 287

## RESULT 8

Q8WZ22 PRELIMINARY; PRT; 1011 AA.  
AC Q8WZ22;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE CAMP-GEFI.  
GN Homo sapiens (Human).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21564207; PubMed=11707077;  
RA Ueno H., Shibasaki T., Iwanaga T., Takahashi K., Yokoyama Y.,  
RA Liu L.M., Yokoi N., Ozaki N., Matsukura S., Yano H., Saito S.;  
RT Characterization of the gene EPAC2: structure, chromosomal  
RT localization, tissue expression, and identification of the liver-  
RT specific isoform.  
RL Genomics 78:91-98(2001).  
DR EMBL: AB027471; BAB72179.1; -  
DR InterPro: IPR000595; cAMP\_binding.  
DR InterPro: IPR000591; DEP.  
DR InterPro: IPR000651; RasGEFN.  
DR InterPro: IPR001895; RasGRF\_CCC25.  
DR Pfam: PF00027; cAMP\_binding; 2.  
DR Pfam: PF00610; DEP; 1.  
DR Pfam: PF00617; RasGEF; 1.  
DR Pfam: PF00618; RasGEFN; 1.  
DR SMART: SM00100; cAMP; 2.  
DR SMART: SM00049; DEP; 1.  
DR SMART: SM00147; RasGEF; 1.  
DR SMART: SM00229; RasGEFN; 1.  
DR PROSITE: PS50042; cAMP\_BINDING\_3; 2.  
DR PROSITE: PS50186; DEP; 1.





```

OY 181 ISRLATHEWRLPHKVRKLYSALERLDPSSMHHRYRLALAKLSPVITPFLDKDMFTI 240
DB 875 VSRRLATWEKLPSEKFKYAEFESIMDSRNHRATRLVAKLEPLIPFMDLLIKDMFT 934
OY 241 HEGNHTLVENLINFEXKMMARARMLHCRSHNPVPLSPRSVSHLHEDSOVARISTC 300
DB 935 HEGNKTFTIDNLVNEFKMMIMANTARTVRYYSQ---PFPDAQAANKHODVR----- 984
OY 301 SEQSLSTRSPASTWAVYVOOLKVIDNORELSRLSRELEP 338
DB 985 -----SYRQOLNVINDNORTLSQMSHRLP 1008

RESULT 11
O9CWP9 PRELIMINARY; PRT; 696 AA.
AC 08VIP9;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE CAM-GEFII (Fragment).
GN CGEF2 OR EPAC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=21564207; PubMed=11707077;
RA Ueno H., Shibasaki T., Iwanaga T., Takahashi K., Yokoyama Y.,
RA Liu L.M., Yokoi N., Ozaki N., Matsukura S., Yano H., Saito S.;
RT "Characterization of the gene EPAC2: structure, chromosomal
RT localization, tissue expression, and identification of the liver-
RT specific isoform."
RL Genomics 78:91-98(2001).
DR EMBL; AB037668; BAB72180.1; -.
DR MGD; MGI:1917723; Cgef2.
DR InterPro; IPR002373; CAMP_kin.
DR InterPro; IPR000595; CNMP_binding.
DR InterPro; IPR000651; RasGEFN.
DR InterPro; IPR001895; RasGRF_CDC25.
DR Pfam; PF00027; CNMP_binding; 1.
DR Pfam; PF00617; RasGEF; 1.
DR Pfam; PF00618; RasGEFN; 1.
DR PRINTS; PR00103; CAMPKINASE.
DR SMART; SM00100; CNMP; 1.
DR SMART; SM00147; RasGEF; 1.
DR SMART; SM00229; RasGEFN; 1.
DR PROSITE; PS50042; CNMP_BINDING_3; 1.
FT NON_TER 696
SQ SEQUENCE 696 AA; 79440 MW; 4006240D4B85A747 CRC64;

Query Match 48.6%; Score 855.5; DB 11; Length 696;
Best Local Similarity 49.7%; Pred. No. 1.9e-69;
Matches 168; Conservative 61; Mismatches 84; Indels 25; Gaps 3;

OY 1 MAALAOEDGWTKGOVLVYVNSAGDAIGLPDARGVATSLGNERLFVYVNPQEVHLLHP 60
DB 381 ISAVADKLGSGGLIIVKMSNGEKVYVLSNDVSVFTLLTNGRLFACPRQFDSLTPLP 440
OY 61 DQLGPTVSABGLDVSADKLAGLTDHWSLFSNISHOVELIHVYLGPOHLDVTTANLE 120
DB 441 EDEGPTTGVTGFFELMSSKDLAYQMTYDMELFNCVHELELLYHTFG-RHNFKKTTANLD 499
OY 121 RMRKRFNELQYVAVAEELICLPVPRQALLRKFTILAAHLKOKUNLSFAPMFGLSNA 180
DB 500 LFLRFNEIQFVWVEVCISQLSKRYDLKFKFIIAAHCKEYKNLNSFAIVMGSNA 559
OY 181 ISRLATHEWRLPHKVRKLYSALERLDPSSMHHRYRLALAKLSPVITPFLDKDMFTI 240
DB 560 VSRRLATWEKLPSEKFKYAEFESIMDSRNHRATRLVAKLEPLIPFMDLLIKDMFT 619

```

```

OY 241 HEGNHTLVENLINFEXKMMARARMLHCRSHNPVPLSPRSVSHLHEDSOVARISTC 300
DB 620 HEGNKTFTIDNLVNEFKMMIMANTARTVRYYSQ---PFPDAQAANKHODVR----- 669
OY 301 SEQSLSTRSPASTWAVYVOOLKVIDNORELSRLSRELEP 338
DB 670 -----SYRQOLNVINDNORTLSQMSHRLP 693

RESULT 12
O9CWP9 PRELIMINARY; PRT; 699 AA.
AC 09CWP9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE 130003D15RIK protein (Fragment).
GN CGEF2 OR 130003D15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Aachii J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batlora S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochika H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,
RA Schimpl L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlaya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyoko-Oka K., Wang K.H., Weitz C., Wittaker C., Wilming L.,
RA Wyszynski B.A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK004874; BAB3633.1; -.
DR MGD; MGI:1917723; Cgef2.
DR InterPro; IPR002373; CAMP_kin.
DR InterPro; IPR000595; CNMP_binding.
DR InterPro; IPR000651; RasGEFN.
DR InterPro; IPR001895; RasGRF_CDC25.
DR Pfam; PF00027; CNMP_binding; 1.
DR Pfam; PF00617; RasGEF; 1.
DR Pfam; PF00618; RasGEFN; 1.
DR PRINTS; PR00103; CAMPKINASE.
DR SMART; SM00100; CNMP; 1.
DR SMART; SM00147; RasGEF; 1.
DR SMART; SM00229; RasGEFN; 1.
DR PROSITE; PS50042; CNMP_BINDING_3; 1.
FT NON_TER 699
SQ SEQUENCE 699 AA; 79799 MW; AD2602AB9A414EA4 CRC64;

Query Match 48.6%; Score 855.5; DB 11; Length 699;
Best Local Similarity 49.7%; Pred. No. 1.9e-69;
Matches 168; Conservative 61; Mismatches 84; Indels 25; Gaps 3;

OY 1 MAALAOEDGWTKGOVLVYVNSAGDAIGLPDARGVATSLGNERLFVYVNPQEVHLLHP 60
DB 384 ISAVADKLGSGGLIIVKMSNGEKVYVLSNDVSVFTLLTNGRLFACPRQFDSLTPLP 443
OY 61 DQLGPTVSABGLDVSADKLAGLTDHWSLFSNISHOVELIHVYLGPOHLDVTTANLE 120

```

```

DB      444  EDEGPTTGVTGTFELMSSKDLAYQMTTYDWELEFNCVHELELYHTFG-RHNFKKTTAND 502
      :| | | | | : :| | | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
OY      121  REMRPFNELQYVAATELCLCPVPGRAQLRKFKITLAHLKEOKNLSNFAVMEGLSNA 180
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      503  LFLRRFNEIOFWVTEVCLSQLSKRVOLKFKFIKIAHCKEYKNLSNFAVMEGLSNA 562
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
OY      181  ISRLAHTEWRLPHKVRKLYSALERLLDPSMNHRYVRLALAKSPVPIPPMLLKDMTFI 240
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      563  VSRLALTWELKPSKFKKFAEFESLMDPSRNHRATRLTAKEPLPIPPMLLIKDMTFI 622
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
OY      241  HEGNHTLVENLINFEMRMARARAHLNHCRSNHPVLSPLRSRVSHLHEDSQVARISTC 300
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      623  HEGNKTFFIDNLVNEFKRMIAANTARTVRYRSG---PFNPDAQAQANKNODVR----- 672
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
OY      301  SEOSLSTRSPASTAYVOOLKVIDNORELSRLSRELEP 338
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      673  -----SYVOLANVIDNORTLSOMSHRLEP 696
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :

```

## RESULT 13

```

O81XK6  PRELIMINARY: PRT: 867 AA.
ID      081XK6  01-MAR-2003 (Tremblrel. 23, Created)
DT      01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT      01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE      Similar to cAMP-regulated guanine nucleotide exchange factor II.
OS      Homo sapiens (Human).
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Brain;
RA      Strausberg R.;
RL      Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL: BC040183; AAH40183.1;
SQ      SEQUENCE 867 AA: 99604 MW: 6612DAC776B792B5 CRC64:

```

```

Query Match      48.6%; Score 855.5; DB 4; Length 867;
Best Local Similarity 49.7%; Pred. No. 2.5e-69;
Matches 168; Conservative 63; Mismatches 82; Indels 25; Gaps 3;

OY      1  MAALAOEDGWTGQVLYVKNVNSAGDAIGLQPDARGVATSLGNRLRVVNPQVNEHLIPR 60
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      552  ISAVNADKLSSGEGCLTIYKMSGGEKVVYLPKPNVSVFTTLTGRLFAACRREOPDSLTPRP 611
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
OY      61  DOLGPTVGSAGEGLDVSADKQLTDHDMSLFNSIHQVELIHYVLGPOHLRBDVTYANDE 120
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      612  EDEGPTTGVTGTFELMSSKDLAYQMTTYDWELEFNCVHELELYHTFG-RHNFKKTTAND 670
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
OY      121  REMRPFNELQYVAATELCLCPVPGRAQLRKFKITLAHLKEOKNLSNFAVMEGLSNA 180
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      671  LFLRRFNEIOFWVTEVCLSQLSKRVOLKFKFIKIAHCKEYKNLSNFAVMEGLSNA 730
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
OY      181  ISRLAHTEWRLPHKVRKLYSALERLLDPSMNHRYVRLALAKSPVPIPPMLLKDMTFI 240
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      731  VSRLALTWELKPSKFKKFAEFESLMDPSRNHRATRLTAKEPLPIPPMLLIKDMTFI 790
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
OY      241  HEGNHTLVENLINFEMRMARARAHLNHCRSNHPVLSPLRSRVSHLHEDSQVARISTC 300
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      791  HEGNKTFFIDNLVNEFKRMIAANTARTVRYRSG---PFNPDAQAQANKNODVR----- 840
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
OY      301  SEOSLSTRSPASTAYVOOLKVIDNORELSRLSRELEP 338
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      841  -----SYVOLANVIDNORTLSOMSHRLEP 864
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :

```

```

RESULT 14
ID      0921P0  PRELIMINARY: PRT: 993 AA.
AC      0921P0;

```

```

DT      01-MAY-1999 (Tremblrel. 10, Created)
DT      01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT      01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE      CAMP-dependent Rap1 guanine-nucleotide exchange factor.
GN      CGEF2 OR 5730402K07RIK.
OS      Mus musculus (Mouse).
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Brain;
RA      Gaudinault G.E., Takaya K., Vale W.W.;
RT      "A brain CAMP-dependent Rap1 guanine-nucleotide exchange factor.";
RL      Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF115480; AAD09132.1;
DR      MGD: MGI:1917723; Ggef2.
DR      InterPro: IPR002373; CAMP_kin.
DR      InterPro: IPR000595; CNMP_binding.
DR      InterPro: IPR000591; DEP.
DR      InterPro: IPR000651; RasGEF.
DR      InterPro: IPR001895; RasGEF_CDC25.
DR      Pfam: PF00027; CNMP_binding; 2.
DR      Pfam: PF00610; DEP; 1.
DR      Pfam: PF00617; RasGEF; 1.
DR      Pfam: PF00618; RasGEF; 1.
DR      PRINTS: PR00103; CAMPKINASE.
DR      SMART: SM00100; CNMP; 2.
DR      SMART: SM00049; DEP; 1.
DR      SMART: SM00147; RasGEF; 1.
DR      SMART: PS00229; RasGEF; 1.
DR      PROSITE: PS50042; CNMP_BINDING_3; 2.
DR      PROSITE: PS50186; DEP; 1.
SQ      SEQUENCE 993 AA: 113488 MW: F503405FD9A9C12F CRC64:

```

```

Query Match      48.6%; Score 855.5; DB 11; Length 993;
Best Local Similarity 49.7%; Pred. No. 3.1e-69;
Matches 168; Conservative 61; Mismatches 84; Indels 25; Gaps 3;

OY      1  MAALAOEDGWTGQVLYVKNVNSAGDAIGLQPDARGVATSLGNRLRVVNPQVNEHLIPR 60
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      678  ISAVNADKLSSGEGCLTIYKMSGGEKVVYLPKPNVSVFTTLTGRLFAACRREOPDSLTPRP 737
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
OY      61  DOLGPTVGSAGEGLDVSADKQLTDHDMSLFNSIHQVELIHYVLGPOHLRBDVTYANDE 120
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      738  EDEGPTTGVTGTFELMSSKDLAYQMTTYDWELEFNCVHELELYHTFG-RHNFKKTTAND 796
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
OY      121  REMRPFNELQYVAATELCLCPVPGRAQLRKFKITLAHLKEOKNLSNFAVMEGLSNA 180
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      797  LFLRRFNEIOFWVTEVCLSQLSKRVOLKFKFIKIAHCKEYKNLSNFAVMEGLSNA 856
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
OY      181  ISRLAHTEWRLPHKVRKLYSALERLLDPSMNHRYVRLALAKSPVPIPPMLLKDMTFI 240
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      857  VSRLALTWELKPSKFKKFAEFESLMDPSRNHRATRLTAKEPLPIPPMLLIKDMTFI 916
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
OY      241  HEGNHTLVENLINFEMRMARARAHLNHCRSNHPVLSPLRSRVSHLHEDSQVARISTC 300
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      917  HEGNKTFFIDNLVNEFKRMIAANTARTVRYRSG---PFNPDAQAQANKNODVR----- 966
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
OY      301  SEOSLSTRSPASTAYVOOLKVIDNORELSRLSRELEP 338
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
DB      967  -----SYVOLANVIDNORTLSOMSHRLEP 990
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :

```

```

RESULT 15
ID      09E0Z6  PRELIMINARY: PRT: 1011 AA.
AC      09E0Z6;
DT      01-MAR-2001 (Tremblrel. 16, Created)
DT      01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT      01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE      CAMP-GEFII.
GN      CGEF2 OR CAMP-GEFII.

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: October 7, 2003, 06:57:32 ; Search time 24 Seconds

(without alignments)  
662,293 Million cell updates/sec

Title: US-09-856-679-2

Perfect score: 1759  
Sequence: 1 MAALQEDGWTGCVLVKVN.....QLKVIDNQRELSRSLSELEP 338

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	566	32.2	1234	1 YNKS_CAEEL	P34578 caenorhabd1
2	283	16.1	1077	1 C3G_HUMAN	Q13905 homo sapien
3	261	14.8	1571	1 C3G_DROME	O77086 drosophila
4	225	12.8	1095	1 CC25_SACKL	O02342 saccharomyc
5	208.5	11.9	911	1 STE6_SCHPO	P26674 schizosacch
6	206	11.7	1244	1 GNRP_RAT	P28818 rattus norv
7	201.5	11.5	1262	1 GNRP_MOUSE	P27671 mus musculu
8	199	11.3	1275	1 GNRP_HUMAN	Q13972 homo sapien
9	196.5	11.2	1373	1 CC25_CANAL	P43069 candida alb
10	193	11.0	1339	1 SOS1_MOUSE	O62245 mus musculu
11	193	11.0	1333	1 SOS1_HUMAN	O07889 homo sapien
12	188.5	10.7	1332	1 SOS2_HUMAN	O07890 homo sapien
13	181	10.3	1252	1 SC25_YEAST	P14771 saccharomyc
14	178	10.1	1589	1 CC25_YEAST	P04821 saccharomyc
15	178	10.1	1595	1 SOS_DROME	P26675 drosophila
16	177.5	10.1	1297	1 SOS2_MOUSE	O02384 mus musculu
17	175.5	10.0	777	1 RGL2_HUMAN	O15211 homo sapien
18	173.5	9.9	778	1 RGL2_MOUSE	O61193 mus musculu
19	169.5	9.6	538	1 BUD5_YEAST	P25300 saccharomyc
20	157.5	9.0	1435	1 LTEL_YEAST	P07866 saccharomyc
21	153.5	8.7	852	1 GND5_MOUSE	O03385 mus musculu
22	150	8.5	914	1 GND5_HUMAN	O12967 homo sapien
23	148.5	8.4	768	1 RGL1_HUMAN	O99216 homo sapien
24	143	8.1	768	1 RGL1_MOUSE	O60655 mus musculu
25	142.5	8.1	895	1 GND5_RAT	O03386 rattus norv
26	108	6.1	814	1 OPHL_HUMAN	O99041 homo sapien
27	94	5.3	910	1 SC15_YEAST	P22224 saccharomyc
28	94	5.3	1844	1 POLR_TYVA	P20128 turnip yell
29	93	5.3	452	1 ETV6_HUMAN	P41212 homo sapien
30	92.5	5.3	1844	1 POLR_TYMV	P10358 turnip yell
31	92	5.2	1594	1 TFC3_YEAST	P46678 saccharomyc
32	92	5.2	687	1 NPH1_MOUSE	O9953 mus musculu
33	91.5	5.2	1844	1 POLR_TYVWC	P28477 turnip yell

34	89.5	5.1	426	1 AGA2_ECOLI	P42903 escherichia
35	89.5	5.1	880	1 SYV_BACST	P11931 bacillus st
36	88.5	5.0	282	1 TAUD_ECOLI	P37610 escherichia
37	88.5	5.0	1321	1 IFX3_DICDI	O15818 dictyostell
38	88.5	5.0	2594	1 7LES_DROVI	P20806 drosophila
39	88	5.0	839	1 Y422_MYCPN	P75175 mycoplasma
40	88	5.0	1165	1 POL_GALV	P21414 gibbon ape
41	87.5	5.0	3321	1 PCN2_HUMAN	O95613 homo sapien
42	87.5	5.0	3674	1 SPCR_HUMAN	O99066 homo sapien
43	87	4.9	376	1 CRPF_RHOSH	P54906 rhodobacter
44	86	4.9	192	1 RR4B_CYACA	O22029 cyanidium c
45	86	4.9	306	1 Y910_TREPA	O83880 treponema p

## ALIGNMENTS

```

RESULT 1
ID YNKS_CAEEL STANDARD: PRT; 1234 AA.
AC P34578:
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein T20G5.5 in chromosome III.
GN T20G5.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN-Bristol N2;
RA Berts M., Smith A.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Durbin R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 DEP domain.
CC -1- SIMILARITY: Contains 2 cyclic nucleotide-binding domains.
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
EMBL: Z30423; CAAB3013.2; -
DR HSSP: P00515; ZBPK.
DR WormRep: T20G5.5; CE23992.
DR InterPro: IPR002373; CAMP_kin.
DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR000591; DEP.
DR InterPro: IPR000651; RasGEFN.
DR InterPro: IPR001895; RasGEF_CDC25.
DR Pfam: PF00027; CNMP_binding; 2.
DR Pfam: PF00610; DEP; 1.
DR Pfam: PF00617; RasGEF; 1.
DR Pfam: PF00618; RasGEFN; 1.
DR PRINTS: PR00103; CAMPKINASE.
DR SMART: SM00100; CNMP; 2.
DR SMART: SM00049; DEP; 1.
DR SMART: SM00147; RasGEF; 1.
DR SMART: SM00229; RasGEFN; 1.
DR PROSITE: PS00888; CNMP_BINDING_1; FALSE_NEG.
DR PROSITE: PS00889; CNMP_BINDING_2; FALSE_NEG.
DR PROSITE: PS50042; CNMP_BINDING_3; 2.
DR PROSITE: PS50186; DEP; 1.

```



```

Db      818 ATSSOPLAARGVA-----ARPGTLHDFSH-----EIAE 846
Oy      84 QLTLDHMSLFNSHVOELLIVGLPOHLRPVTANLEREMRRELOLYWATELCLPVP 143
Db      847 QLTLLDAELF---YKTEIPVLLMAKQNEKSPKSNLOFHEHNKMSYWRISIMIOEKA 903
Oy      144 GPRAQLLRKFKIKLAHLKEQKNLNSPFAVMFGLSNSAISRLAHTWELPKVRKLYSALE 203
Db      904 QDBERLLKFKIKIMKHLKRLNNSYALISALDSAIRL--EMOKOTSEGLAEYCT-- 959
Oy      204 RLDDPSMNRHVRLALAKLSPPVLPMPPLLLKDMFTIEGNNHTLVENLNFEEK----- 256
Db      960 -LIDSSSFRAVRAALSEVEPPCIPYLGILIDLTFEVHLNPNPYIDCKVNFESKRMQOFNI 1018
Oy      257 ---MRMARAAKMLHCHRSNHPVPLSRVSHLHDS 292
Db      1019 LDSMRCPQQA---HYDMRRNDITINFNDPSHLAEBA 1053

RESULT 3
C3G_DROME STANDARD: PRT: 1571 AA.
ID_C3G_DROME
AC 077086; Q9W3W3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Guanine nucleotide-releasing factor 2 (CRK SH3-binding GNRP).
GN C3G OR CG3136.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
OX
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM SHORT), AND CHARACTERIZATION.
RC TISSUE=Eye imaginal disk;
RX MEDLINE=99094900; Pubmed=9878058;
RA Ishihara S., Gaul U., Hanafusa H.;
RT "Activation of the Drosophila C3G leads to cell fate changes and
RT overproliferation during development, mediated by the RAS-MAPK
RT pathway and RAPI.";
RL EMBO J. 18:145-155(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang O., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayan A., An H.-U., Andrews-Pinnock C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokov D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahke C., Davenport L.B., Davies P.,
RA De Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,
RA Folsler C., Gabriellani A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin L., Houston K.A., Howard T.J., Wei M.-H., Ibegwam C.,
RA Jallat M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,

```

```

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [3]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RC STRAIN=Berkley;
RX MEDLINE=22426069; Pubmed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [4]
RP SEQUENCE OF 1036-1571 FROM N.A.
RC STRAIN=Berkley; TISSUE=Ovary;
RX MEDLINE=22426066; Pubmed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarini H., Kronmiller B., Paclob J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -1- FUNCTION: Guanine nucleotide-releasing protein that binds to SH3
CC domain of CRK. Transduces signals from CRK to activate RAS. Also
CC involved in MAPK activation.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=077086-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=077086-2; Sequence=VSP_001823;
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- DEVELOPMENTAL STAGE: Throughout development.
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 39.
CC -1- CAUTION: Ref.4 sequence differs from that shown due to a
CC frameshift in position 1366.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AF053338; AAC35280.1; ALT_FRAME.
CC EMBL: AE003458; AAF46200.2;
CC EMBL: AY113355; AAM29360.1; ALT_FRAME.
CC FLYBase: FBgn0026145; C3G.
CC GO: GO:0005088; F:RAS; guany1-nucleotide exchange factor activity; IMP.
CC GO: GO:0007265; P:RAS protein signal transduction; IMP.
CC InterPro: IPR000651; RASGEFN.
CC InterPro: IPR001895; RASGRF_CDC25.
CC Pfam: PF00617; RASGEF_1.
CC Pfam: PF00618; RASGEFN_1.
CC SMART: SM00147; RASGEF.
CC SMART: SM00229; RASGEFN_1.

```

DR PROSITE; PS00720; RASGEF; 1.  
 DR PROSITE; PS50009; RASGEF\_CAT; 1.  
 DR PROSITE; PS50212; RASGEF\_NTER; 1.  
 KW Guanine-nucleotide releasing factor; SH3-binding;  
 Developmental protein; Alternative splicing.  
 FT DOMAIN 49  
 FT DOMAIN 186  
 FT DOMAIN 195  
 FT DOMAIN 198  
 FT DOMAIN 213  
 FT DOMAIN 216  
 FT DOMAIN 238  
 FT DOMAIN 247  
 FT DOMAIN 639  
 FT DOMAIN 644  
 FT DOMAIN 699  
 FT DOMAIN 702  
 FT DOMAIN 735  
 FT DOMAIN 787  
 FT DOMAIN 791  
 FT DOMAIN 806  
 FT DOMAIN 814  
 FT DOMAIN 815  
 FT DOMAIN 818  
 FT DOMAIN 1170  
 FT DOMAIN 1292  
 FT DOMAIN 1462  
 FT DOMAIN 1465  
 FT SITE 546  
 FT SITE 556  
 FT SITE 820  
 FT SITE 831  
 FT SITE 924  
 FT SITE 935  
 FT SITE 986  
 FT SITE 997  
 FT VARSPLIC 668  
 FT 718  
 FT 19  
 FT 93  
 FT 154  
 FT 202  
 FT 202  
 FT 243  
 FT 243  
 FT 450  
 FT 453  
 FT 477  
 FT 477  
 FT 807  
 FT 818  
 FT 851  
 FT 1004  
 FT 1092  
 FT 1127  
 FT 1127  
 FT 1311  
 FT 1311  
 SQ SEQUENCE 1571 AA; 172049 MW; B9EF9E8670457684 CRC64;  
 Query Match 14.8%; Score 261; DB 1; Length 1571;  
 Best Local Similarity 33.3%; Pred. No. 4.5e-14;  
 Matches 67; Conservative 43; Mismatches 73; Indels 18; Gaps 7;  
 QY 64 GPTVGSAGC-----LDIVSAKDLAGQLTDHDSLFNSIHOVELIHYVVGPOHLRDVTTA 117  
 DB 1321 GGCSTAGAGGQNPSSLDLKLST-ELAEQMTLLDAELFT--KLEIPEVLFFAKDQCEEKSP 1376  
 QY 118 NLERMRKRFNELQYVAVATELCIPVGPRAQLRKFEIKLAHLKQKNLNSFEAVMFGLS 177  
 DB 1377 NLRKTEHFNRKSWYARSKILRLDPAKERKRVNFEIKIMKHLRMMNNYSTLLALLSALD 1436  
 QY 178 NSAISRLAHWTWR-LPHKVRKLYSALERLDPNMNHRVYRLALAKLSPVPIPFMDLLKD 236  
 DB 1437 SGPIRL--ENQKGTTEVRSCA-----LIDSSSFRAVROALATNPICPIYIGLLIOD 1490  
 QY 237 MTFHEGNT--LVENLINFEEK 256  
 DB 1491 LTFVHVGNDYLSKGVINFSK 1511  
 RESULT 4  
 CC25\_SACKL STANDARD; PRT; 1095 AA.  
 AC 002342;  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Cell division control protein 25 (Fragment).  
 GN CDC25.  
 OS Saccharomyces kluyveri (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI\_TaxID=4934;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=92354938; PubMed=1644315;  
 RA Prigozy T., Gonzales E., Broek D.;  
 RT "Identification and analysis of a DNA fragment from Saccharomyces  
 kluyveri that can complement the loss of CDC25 function in  
 Saccharomyces cerevisiae."  
 RT Gene 117:67-72(1992).  
 RL Gene 117:67-72(1992).  
 CC -! FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP. THIS  
 CC PROTEIN POSITIVELY CONTROLS THE LEVEL OF CELLULAR CAMP AT START,  
 CC THE STAGE AT WHICH THE YEAST CELL DIVISION CYCLE IS TRIGGERED.  
 CC -! SIMILARITY: Contains 1 Ras-GEF domain.  
 CC -! SIMILARITY: Contains 1 N-terminal Ras-GEF domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M82964; AAA34479.1; -  
 DR PIR: PC1114; PC1114.  
 DR InterPro: IPR000651; RASGEF.  
 DR InterPro: IPR001895; RASGEF\_CDC25.  
 DR Pfam: PF00617; RASGEF; 1.  
 DR Pfam: PF00618; RASGEF; 1.  
 DR SMART: SM00147; RASGEF; 1.  
 DR SMART: SM00229; RASGEF; 1.  
 DR PROSITE; PS00720; RASGEF; 1.  
 DR PROSITE; PS50009; RASGEF\_CAT; 1.  
 DR PROSITE; PS50212; RASGEF\_NTER; 1.  
 KW Guanine-nucleotide releasing factor; Cell division; Cell cycle;  
 KW Mitosis; Transmembrane.  
 FT NON\_TER 1  
 FT TRANSMEM 959  
 FT DOMAIN 626  
 FT DOMAIN 812  
 FT SEQUENCE 1095 AA; 125605 MW; 7A8B2F90E31A44AC CRC64;  
 Query Match 12.8%; Score 225; DB 1; Length 1095;  
 Best Local Similarity 27.4%; Pred. No. 3.4e-11;  
 Matches 75; Conservative 44; Mismatches 131; Indels 24; Gaps 7;  
 QY 12 KGOVLVK--VNSADAIIGLPDARGAVATSLGLNRLTFVNPQEVNHLIPHPDOLGPTVG 68  
 DB 752 KGRLSMKGNLKNFVESINFSDGSSSTTV-----POSSRSSVSAVGVSSSTTG 800  
 QY 69 -----SABGLDVSANDLAGQLTDHDSLFNSIHOVELIHYVVGPOHLRDVTTANLFRMR 124  
 DB 801 FRMRKLKLD-IDSLDYAKOLTIKHSLEFKISPECIDRTWGNKNCNMGSKNTEFTIS 859  
 QY 125 RFNELQYVAVATELCIPVGPRAQLRKFEIKLAHLKQKNLNSFEAVMFGLSAISRL 184  
 DB 860 NSNHLITNVSEFVAVQTDIKRIQIQLQFFINVAANCHLNNSFTALISALYSSPIRL 919  
 QY 185 AHTWERLPHKVRKLYSALERLDPNMNHRVYRLALAKLSP-VPIPFMDLLKDMTFHEG 243  
 DB 920 KRTMAAVPEYKKLLEELNLTMDSAKNFLRYROLKLSIGDFPCVFFSVYLSDLTFANG 979  
 QY 244 NHTLVEN--LINF-EKKRMARARARMLHHCRSH 273  
 DB 980 NPDFLHRTVLVNFGRVRIILEIKELISVYORSH 1013  
 RESULT 5  
 STE6\_SCHPO STANDARD; PRT; 911 AA.  
 AC P26674; Q9JUM8;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)



DT 15-SEP-2003 (Rel. 42, last annotation update)  
DE Sleg protein.  
GN SLEG OR SPC1442.01 OR SPC1450.17.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90190870; PubMed=2107403;  
RA Hughes D.A., Fukui Y., Yamamoto M.;  
RT "Homologous activators of ras in fission and budding yeast.";  
RL Nature 344:355-357(1990).  
[2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=972;  
RC MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
RA Gentles S., Godle A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skellon J., Slimmons M., Squares R., Squares S., Stevens K.,  
RA Taylor R., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,  
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerretti L., Lowe T., McCormick W.R., Paulsen I., Potashkin J.,  
RA Shparovskii G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP. THIS  
CC PROTEIN IS ESSENTIAL FOR MATING.  
CC -1- SIMILARITY: Contains 1 SH3 domain.  
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.  
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X53254; CAJ37345.1; -;  
DR EMBL: AL049559; CAB40184.1; -;  
DR EMBL: AL031966; CAJ21435.1; -;  
DR PIR: S28098; S28098.  
DR GeneDB Spombe: SPC1442.01; -;  
DR InterPro: IPR000651; RasGEF.  
DR InterPro: IPR001895; RasGEF\_CDC25.  
DR InterPro: IPR001452; SH3.  
DR Pfam: PF00617; RasGEF\_1.  
DR Pfam: PF00618; RasGEF\_N\_1.  
DR Pfam: PF00018; SH3\_1.  
DR SMART: SM00147; RasGEF\_1.  
DR SMART: SM00229; RasGEF\_N\_1.  
DR SMART: SM00326; SH3\_1.  
DR PROSITE: PS00720; RasGEF\_1.

DR PROSITE: PS50009; RasGEF\_CAT; 1.  
DR PROSITE: PS50212; RasGEF\_NTER; 1.  
DR PROSITE: PS50002; SH3\_1.  
KW Guanine-nucleotide releasing factor; SH3 domain.  
FT DOMAIN 1 60 SH3.  
FT DOMAIN 489 623 N-TERMINAL RAS-GEF.  
FT DOMAIN 663 894 RAS-GEF.  
SQ SEQUENCE 911 AA; 105185 MW; 5F8C12020C4B753F CRC64;  
Query Match 11.98; Score 208.5; DB 1; Length 911;  
Best Local Similarity 26.28; Pred. No. 6; Be-10;  
Matches 63; Conservative 43; Mismatches 107; Indels 27; Gaps 5;  
QY 51 QEVHELIPH--PQGLGTVASAGLDVSAKLAGQUTDHDWSLFNSIHQVELIHVYLG 108  
Db 638 QLEEDLSLHNSPD---PIIKDELIVLLPRETRAKQCLIEFQSFISITSIQFLTKWD- 693  
QY 109 QHLRDVTTANLIERFMRR-----FNELOYVWATELCLCPVGPRAQLRKFKLAHL 160  
Db 694 -----NLNFSPEKSTFYLSNHLVNFYETETIVQEEPRRRRTNVLAVFYQCDYL 744  
QY 161 KEQKUNSFPAVWFGLSNSAISRLAHTWELPKVKRLYSALERLDPSSNNHRYRLALA 220  
Db 745 RELNNFASLEFSIISALNSSPIHRLRTWAMLNSKTLASFLLNNLTGARRKFSNYRDCLE 804  
QY 221 KLSPPVYFPMPLLLKQWTFHEGNNHTLVENLNEPKRMMAARAARMLHGRS---HNPV 276  
Db 805 NCVLPVPLGVFTFTDLTKGNKDFQNMINEKTKVTRILNETIKKFSQGYMENPI 864  
RESULT 6  
GNRP\_RAT  
ID GNRP\_RAT STANDARD; PRT; 1244 AA.  
AC P28818;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Guanine nucleotide releasing protein (GNRP) (P140 Ras-GRF).  
GN RASGR1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=92350260; PubMed=1379346;  
RT Shou C., Farusworth C.L., Neel B.G., Feig L.A.;  
RT "Molecular cloning of cDNAs encoding a guanine-nucleotide-releasing  
RT factor for Ras p21.";  
RL Nature 358:351-354(1992).  
CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP.  
CC -1- SIMILARITY: Contains 1 DBL-homology (DH) domain.  
CC -1- SIMILARITY: Contains 2 PH domains.  
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.  
CC -1- SIMILARITY: Contains 1 IO domain.  
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X67241; CAA47666.1; -;  
DR PIR: S29083; S29083.  
DR InterPro: IPR001331; GDS\_CDC24.  
DR InterPro: IPR000048; IQ\_region.  
DR InterPro: IPR001849; PH.  
DR InterPro: IPR000651; RasGEF.  
DR InterPro: IPR001895; RasGEF\_CDC25.

```
DR InterPro: IPR000219; RhoGEF.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00169; PH; 2.
DR Pfam: PF00617; RasGEF; 1.
DR Pfam: PF00618; RasGEFN; 1.
DR Pfam: PF00621; RhoGEF; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00235; PH; 2.
DR SMART: SM00147; RasGEF; 1.
DR SMART: SM00229; RasGEFN; 2.
DR SMART: SM00325; RhoGEF; 1.
DR PROSITE: PS00741; DH_1; 1.
DR PROSITE: PS50010; DH_2; 1.
DR PROSITE: PS00720; RasGEF; 1.
DR PROSITE: PS50096; IQ; 1.
DR PROSITE: PS50003; PH_DOMAIN; 2.
DR PROSITE: PS50009; RasGEF_CAT; 1.
DR PROSITE: PS50212; RasGEF_NTER; 1.
KM Guanine-nucleotide releasing factor; Repeat.
FT DOMAIN 22 129 PH 1.
FT DOMAIN 204 229 IQ.
FT DOMAIN 240 426 DH.
FT DOMAIN 456 582 PH 2.
FT DOMAIN 629 743 N-TERMINAL RAS-GEF.
FT DOMAIN 1009 1241 RAS-GEF.
SQ SEQUENCE 1244 AA; 142666 MW; 4B647879E842AF6B CRC64;

Query Match 11.7%; Score 206; DB 1; Length 1244;
Best Local Similarity 23.9%; Pred. No. 1.7e-09;
Matches 68; Conservative 56; Mismatches 144; Indels 16; Gaps 6;

QY 28 LQPDARVATSL--GLNRLVYVNPQEVHELPHPDQGLPVYGSAGGLDLYSAKLAQL 85
DB 963 LIPQKRAAAMIMRLTQEEITENSMDELTLMEGV----KTEPPENHSAMEIAQL 1017
QY 86 TDHWSLNSIHQVELIHVYGPQHL--RDVTANLRPMKRFVELGYWATEICLPV 142
DB 1018 TLDDLVKSLPYEEF---FGGGMKADKNERTPIYIKTRHFHISNLSASELRDE 1073
QY 143 PGPRALRKFTIKLAHLKEQKLNLSFPAVMEGLNSAISRLAHTWERLPHKRVLYSAL 202
DB 1074 VSARASTIEKWWAVADICRCLHNVALEITSSIRSAIFRLKKTWLVSKOTKSLPDKL 1133
QY 203 ERLDPSNNHRYRLAKLSPVYIPMPYLKMTFTHEGHTLVEN-LINFEKRMMA 261
DB 1134 QLVSSDGRFKMLRFTLNCDDPCVPYLGMYLTDLAFLEEGTPNTEGLVNFSKRMIS 1193
QY 262 RAARMLHCRSHNPVPLSLRSRVSHLHEDSOVARISTCSQSL 305
DB 1194 HIRREIROFO-QTYKIEPQKRVGYLVDEFVLDESLYEASL 1236

RESULT 7
GNRP_MOUSE STANDARD; PRT; 1262 AA.
AC P27671;
DT 01-AUG-1992 (rel. 23, Created)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 15-SEP-2003 (rel. 42, Last annotation update)
DE Guanine nucleotide releasing protein (GNRP) (Ras-specific nucleotide
DE exchange factor CDC25) (CDC25Mm).
DE RASGRF1 OR CDC25 OR GRF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAUB/c;
RX MEDLINE=93010996; PubMed=1396590;
RA Cen H., Lowy D.D.;
RT "Isolation of multiple mouse cDNAs with coding homology to
RT Saccharomyces cerevisiae CDC25. Identification of a region related to
```

```
RT Bcr, Vav, Dbl and CDC24.";
RN EMBL J. 11:4007-4015(1992).
RN [2]
RP SEQUENCE OF 791-1262 FROM N.A.
RP STRAIN=Swiss; TISSUE=Brain;
RC MEDLINE=92289680; PubMed=1376246;
RX Martegani E., Vanoni M., Zippel R., Coccetti P., Brambilla R.,
RA Ferrari C., Scutari E.P., Alberghina L.,
RT "Cloning by functional complementation of a mouse cDNA encoding a
RT homologue of CDC25, a Saccharomyces cerevisiae Ras activator.";
RN EMBL J. 11:2151-2157(1992).
RN [3]
RP SEQUENCE OF 1031-1226 FROM N.A.
RX MEDLINE=92357779; PubMed=1379731;
RA Wel W., Mosteller R.D., Sanyal P., Gonzales E., McKinney D.,
RA Dasgupta C., Li P., Liu B.X., Broek D.;
RT "Identification of a mammalian gene structurally and functionally
RT related to the CDC25 gene of Saccharomyces cerevisiae.";
RN Proc. Natl. Acad. Sci. U.S.A. 89:7100-7104(1992).
RL - FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP.
CC - TISSUE SPECIFICITY: BRAIN.
CC - SIMILARITY: Contains 2 PH domains.
CC - SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC - SIMILARITY: Contains 1 Ras-GEF domain.
CC - SIMILARITY: Contains 1 IQ domain.
CC - SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L20899; AAA02741.1; -.
DR EMBL: X59868; CAA42525.1; -.
DR PIR: S28407; S28407.
DR MGI: 99694; Rasgrf1.
DR InterPro: IPR001331; GDS_CDC24.
DR InterPro: IPR000048; IQ_region.
DR InterPro: IPR001849; PH.
DR InterPro: IPR000651; RasGEFN.
DR InterPro: IPR001895; RasGRE_CDC25.
DR InterPro: IPR000219; RhoGEF.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00617; RasGEF; 1.
DR Pfam: PF00618; RasGEFN; 1.
DR Pfam: PF00621; RhoGEF; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00235; PH; 2.
DR SMART: SM00147; RasGEF; 1.
DR SMART: SM00229; RasGEFN; 2.
DR SMART: SM00325; RhoGEF; 1.
DR PROSITE: PS00741; DH_1; 1.
DR PROSITE: PS50010; DH_2; 1.
DR PROSITE: PS00720; RasGEF; 1.
DR PROSITE: PS50096; IQ; 1.
DR PROSITE: PS50003; PH_DOMAIN; 2.
DR PROSITE: PS50009; RasGEF_CAT; 1.
DR PROSITE: PS50212; RasGEF_NTER; 1.
KM Guanine-nucleotide releasing factor; Repeat.
FT DOMAIN 22 130 PH 1.
FT DOMAIN 208 233 IQ.
FT DOMAIN 244 430 DH.
FT DOMAIN 460 588 PH 2.
FT DOMAIN 635 749 N-TERMINAL RAS-GEF.
FT DOMAIN 1027 1259 RAS-GEF.
FT DOMAIN 1033 1033 E-> D (IN REF. 3).
SQ SEQUENCE 1262 AA; 144101 MW; 38BE68F7C228DC8 CRC64;

Query Match 11.5%; Score 201.5; DB 1; Length 1262;
```

Best Local Similarity 25.1%; Pred. No. 4,2e-09;  
Matches 58; Conservative 51; Mismatches 115; Indels 7; Gaps 4;

OY 78 AKDLAQGLTDHMSLSPNSIHQVELI--HYVLGPOHLRDVTTANLERMRBRENELOVAVAT 135  
Db 1028 ALAIEADLTLDHLVFSEIEEFGGWMKAEEYER--TPYIMKTKTHFNHVSNTIAS 1084  
OY 136 ELCLCPGPRPAQLLRKFIKLAHLKEQKNLNSFFAVMFLSNSAISRLAHTWERLPHKV 195  
Db 1085 EIIRNEDISARASAEKVAIVADICRCLHNYNAVLETTSSINSALFRLKTKMLKVSQOT 1144  
OY 196 RKLYSALERLLDPSWNRHYRLALAKISPVYIPMPILLKDMFIEHCNHTLVN-LINF 254  
Db 1145 KSLDLKLOKLVSSEGRFKNRESLRNCDPCVPYLGMYLTDLVFEIEGTPTNYTEDGLVNF 1204  
OY 255 EKRMARARARMLHHCHSHNPVPLSPRSRVSHLHESQVARISTCEOSL 305  
Db 1205 SKRMMSHIREIROFO-OTTYKIDPOPKVIOYLDESFMLDESLYESSL 1254

RESULT 8  
GNRP\_HUMAN STANDARD; PRT: 1275 AA.

AC 013972;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Guanine nucleotide releasing protein (GNRP) (Ras-specific nucleotide exchange factor CDC25).  
GN RASGRF1 OR CDC25.  
OS Homo sapiens (Human); Chordata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95129875; PubMed=7828890;  
RA Wei W., Das B., Park W., Brock D.;  
RT "Cloning and analysis of human CDNas encoding a 140-kDa brain guanine nucleotide-exchange factor, Cdc25GEF, which regulates the function of Ras";  
RT Ras";  
RL Gene 151,279-284(1994).  
CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP.  
CC -1- SIMILARITY: Contains 2 PH domains.  
CC -1- SIMILARITY: Contains 1 DBL-homology (DH) domain.  
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.  
CC -1- SIMILARITY: Contains 1 IO domain.  
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: L26584; AAA58417.1; -;  
CC PIR: A38985; A38985.  
CC Genbank: KCNC:9875; RASGRF1.  
CC GKI: Q13972; -;  
CC GO: GO:0005088; F:RAS guanyl-nucleotide exchange factor activity; TAS.  
CC InterPro: IPR001331; GDS\_CDC24.  
CC InterPro: IPR000048; IO\_region.  
CC InterPro: IPR001849; PH.  
CC InterPro: IPR000651; RASGEFN.  
CC InterPro: IPR001895; RASGEF\_CDC25.  
CC InterPro: IPR000219; RHOGEF.  
CC Pfam: PF00612; IO; 1.  
CC Pfam: PF00169; PH; 2.  
CC Pfam: PF00617; RASGEF; 1.  
CC Pfam: PF00618; RASGEFN; 1.  
CC Pfam: PF00621; RHOGEF; 1.

DR SMART: SM00015; IO; 1.  
DR SMART: SM00233; PH; 2.  
DR SMART: SM00147; RASGEF; 1.  
DR SMART: SM00329; RASGEFN; 2.  
DR SMART: SM00325; RHOGEF; 1.  
DR PROSITE: PS00741; DH\_1; 1.  
DR PROSITE: PS00741; DH\_2; 1.  
DR PROSITE: PS00720; RASGEF; 1.  
DR PROSITE: PS50093; PH\_DOMAIN; 1.  
DR PROSITE: PS50093; PH\_DOMAIN; 2.  
DR PROSITE: PS50009; RASGEF\_CAT; 1.  
DR PROSITE: PS50212; RASGEF\_NTER; 1.  
KW Guanine-nucleotide releasing factor; Repeat.  
FT DOMAIN 22 129  
FT DOMAIN 204 229  
FT DOMAIN 240 426  
FT DOMAIN 467 584  
FT DOMAIN 644 762  
FT DOMAIN 1040 1272  
FT SEQUENCE 1275 AA; 145381 MW; 86C6F54A1E451F1 CRC64;

Query Match 11.3%; Score 199; DB 1; Length 1275;  
Best Local Similarity 27.8%; Pred. No. 7,1e-09;  
Matches 55; Conservative 41; Mismatches 98; Indels 4; Gaps 3;

OY 70 AEGIDLVSAKDLAGQLTDHMSLSPNSIHQVELIHYVLPOLH-RDVTANLERMRBRENE 128  
Db 1033 AEPENISALEIAEQTLTDHLVFKKIPYEEF--FGGWMKLEENETPYIMKTKTHFN 1090  
OY 129 LQYVATELCIPYGPRAQLLRKFIKLAHLKEQKNLNSFFAVMFLSNSAISRLAHTW 188  
Db 1091 ISNLIAEIIIRNEDINRVAIEKVAIVADICRCLHNYNAVLETTSSMNSAISRLKKT 1150  
OY 189 ERLPHKRLKLYSALERLLDPSWNRHYRLALAKISPVYIPMPILLKDMFIEHCNHTLV 248  
Db 1151 LKYSKOTKALIDKLQKLVSSBGRFKNLREALKNCDPCVPYLGMYLTDLAIEEGTPT 1210  
OY 249 EN-LINEKRMARAR 265  
Db 1211 EDGLVNSKRMMSHIRE 1228

RESULT 9  
CC25\_CANAL STANDARD; PRT: 1333 AA.

AC CC25\_CANAL  
ID P43069;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Cell division control protein 25.  
GN CDC25 OR CSC25.  
OS Candida albicans (Yeast).  
CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
CC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
OX NCBI\_TaxID=5476;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=792-1;  
RX MEDLINE=93238685; PubMed=8477693;  
RA Goldberg D., Marbach I., Gross E., Levitzki A., Simchen G.;  
RT "A Candida albicans homolog of CDC25 is functional in Saccharomyces cerevisiae";  
RL Eur. J. Biochem. 213:195-204(1993).  
CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP. THIS PROTEIN POSITIVELY CONTROLS THE LEVEL OF CELLULAR CAMP AT START, THE STAGE AT WHICH THE YEAST CELL DIVISION CYCLE IS TRIGGERED.  
CC -1- SIMILARITY: Contains 1 SH3 domain.  
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.  
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M94160; AAA34329.2; -  
DR HSSP; P29354; IGCF.  
DR InterPro; IPR000651; RASGERN.  
DR InterPro; IPR001895; RASGR\_CDC25.  
DR InterPro; IPR001452; SH3.  
DR Pfam; PF00617; RASGEF. 1.  
DR Pfam; PF00618; RASGEFN. 1.  
DR Pfam; PF00018; SH3. 1.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR PRODOM; PD000066; SH3. 1.  
DR SMART; SM00147; RASGEF. 1.  
DR SMART; SM00229; RASGEFN. 1.  
DR SMART; SM00326; SH3. 1.  
DR PROSITE; PS00720; RASGEF. 1.  
DR PROSITE; PS50009; RASGEF\_CAT. 1.  
DR PROSITE; PS50212; RASGEF\_NTER. 1.  
DR PROSITE; PS50002; SH3. 1.  
KW Guanine-nucleotide releasing factor; Cell division; Cell cycle;  
KW Mitosis; SH3 domain.  
FT DOMAIN 30 SH3.  
FT DOMAIN 900 1034 N-TERMINAL RAS-GEF.  
FT DOMAIN 1068 1305 RAS-GEF.  
SQ SEQUENCE 1333 AA; 152012 MW; DAF690D6AC192483 CRC64;

Query Match 11.2%; Score 196.5; DB 1; Length 1333;  
Best Local Similarity 27.9%; Pred. 1.2e-08;  
Matches 63; Conservative 45; Mismatches 99; Indels 19; Gaps 6;

QY 43 ERLFVNPOEYHELPHPDQLP-TVGSA-----EGLDVSASDIAGOLTDHMSLNF 94  
DB 1033 ERIYV---ERDPVLPNKPAPPLTKGSSLSKKPRVMDIV--ELARQLTRPFLKLYC 1085  
QY 95 SIHOVELIHYVLGPOHLDVTTANIERFMRNRELQYVAIVELCLCPGPRRAQLLRFI 154  
DB 1086 KITKFACLAKYWGKSGLSSESIDSTIPRIKASNQLTNFVGVIIRKADPKRVOILRFEI 1145  
QY 155 KLAHLKQKNLNSFEAFMGISNLSAISRLAHTWRLPHVKRYKLSALERLIDPSMNRV 214  
DB 1146 QVADKCRQYNNFSSMTAIIISLYSPTRILKKTWYIMADALSNLKNMKNLMSNRNNE 1205  
QY 215 YRLALAKL-SPPVPIPEMLLKDMFTIEGNTLVEN--LINFEEK 256  
DB 1206 YRDVLKFTGSEPCVCFEGVYISDLTFVYHGNDPLYNFRQVNFPAK 1251

RESULT 10  
SOS1\_MOUSE STANDARD: PRT; 1319 AA.  
ID SOS1\_MOUSE 062245; 062244; Rel. 38. Created  
AC 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DE 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Son of sevenless protein homolog 1 (SOS-1) (msos-1).  
GN SOS1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Swiss; Tissue=Eye;  
RX MEDLINE=9235328; PubMed=1631150;  
RA Bowtell D., Fu P., Simon M., Senior P.,  
RT "Identification of murine homologues of the Drosophila son of  
RT sevenless gene: potential activators of ras."  
RL Proc. Natl. Acad. Sci. U.S.A. 89:6511-6515(1992).  
RN [2]  
RP STRUCTURE BY NMR OF 415-548.

RX MEDLINE=97360234; PubMed=9217262;  
RA Koshida S., Kigawa T., Kim J.-H., Shirouzu M., Bowtell D.,  
RA Yokoyama S.,  
RT "The solution structure of the pleckstrin homology domain of mouse  
RT Son-of-sevenless 1 (msos1)."  
RL J. Mol. Biol. 269:579-591(1997).  
CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP (BY  
CC SIMILARITY).  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MOST EMBRYONIC AND ADULT TISSUES.  
CC -1- SIMILARITY: Contains 1 DBL-homology (DB) domain.  
CC -1- SIMILARITY: Contains 1 PH domain.  
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.  
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; Z11574; CAAT7662.1; -  
DR EMBL; Z11578; CAAT7665.1; -  
DR PDB; 1PMS; 15-MAY-97.  
DR PDB; 1GBQ; 04-SEP-97.  
DR PDB; 2GBQ; 04-SEP-97.  
DR PDB; 3GBQ; 04-SEP-97.  
DR PDB; 4GBQ; 04-SEP-97.  
DR MGD; MGI:98354; Sos1.  
DR InterPro; IPR001331; GDS\_CDC24.  
DR InterPro; IPR004822; Histone\_core.  
DR InterPro; IPR001849; PH.  
DR InterPro; IPR000651; RASGERN.  
DR InterPro; IPR001895; RASGR\_CDC25.  
DR InterPro; IPR00219; RhoGEF.  
DR Pfam; PF00169; PH. 1.  
DR Pfam; PF00617; RASGEF. 1.  
DR Pfam; PF00618; RASGEFN. 1.  
DR Pfam; PF00621; RhoGEF. 1.  
DR Pfam; PF00233; PH. 1.  
DR SMART; SM00147; RASGEF. 1.  
DR SMART; SM00229; RASGEFN. 1.  
DR SMART; SM00325; RhoGEF. 1.  
DR PROSITE; PS00741; DH. 1; FALSE\_NEG.  
DR PROSITE; PS50010; DH. 2; 1.  
DR PROSITE; PS00720; RASGEF. 1.  
DR PROSITE; PS50003; PH\_DOMAIN. 1.  
DR PROSITE; PS50009; RASGEF\_CAT. 1.  
DR PROSITE; PS50212; RASGEF\_NTER. 1.  
KW Guanine-nucleotide releasing factor; 3D-structure.  
FT DOMAIN 200 DH.  
FT DOMAIN 444 548 PH.  
FT DOMAIN 597 741 N-TERMINAL RAS-GEF.  
FT DOMAIN 780 1019 RAS-GEF.  
FT DOMAIN 1244 1247 POLY-PRO.  
FT TURN 416 418  
FT HELIX 419 427  
FT TURN 443 445  
FT STRAND 446 451  
FT STRAND 459 463  
FT STRAND 467 471  
FT STRAND 490 495  
FT STRAND 500 502  
FT STRAND 507 507  
FT TURN 508 509  
FT STRAND 510 510  
FT STRAND 514 517  
FT STRAND 523 527  
FT HELIX 532 545  
FT TURN 546 546  
SQ SEQUENCE 1319 AA; 150882 MW; 3286088A5BA04A6 CRC64;

Query Match 11.0%; Score 193; DB 1; Length 1319;  
 Best Local Similarity 27.9%; Pred. No. 2.4e-08;  
 Matches 60; Conservative 46; Mismatches 83; Indels 26; Gaps 9;

OY 65 PIV-----GSAGCDLVSAK--DLAQLTDHDSLSPNSHOVELIHYLGPOHLRDT 115  
 DB 759 PIVEMHISRGHIEITDLDLHPIETARQITLESIDLYRAVQSELSGWTRED-KEIN 817  
 OY 116 TAVLEPFRNFNLQYVAVNELCIPVPS--PRAOLRKFIKLAHLKEQKINSPFAVM 173  
 DB 818 SPNLKMRHTTTLTLMF--EKCTIVETNLEERVAVSRILEIQVQELNENGYLEV 875  
 OY 174 FGLSNSAISRLATHTWRLPHKVKLYLSALERLDPSNHRVYRLAKL--SPVPIPFM 230  
 DB 876 SAMNSPVYKRLDTFTFQIDPSROKKI---LEEAEHLSGDH--YKKYLAKLSINPCVPFF 930  
 OY 231 PLLKDMTFIEGHNHTLV---ENLNFEMKRMMA 261  
 DB 931 GIVLTNLKTEGDNPEVLRHKGKELINFSKRRVA 965

RESULT 11  
 SOST\_HUMAN STANDARD; PRT; 1333 AA.  
 AC 007889;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Son of sevenless protein homolog 1 (SOS-1).  
 GN SOS1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Plimates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA MEDLINE=93262494; PubMed=8493579;  
 RA Chardin P., Camonels J.H., Gale N.W., van Aelst L., Wigler M.H.,  
 RA Bar-Sagi D.;  
 RT "Human Sos1: a guanine nucleotide exchange factor for Ras that binds  
 RT to GRB2.";  
 RL Science 260:1338-1343(1993).  
 RN [2]  
 RP STRUCTURE BY NMR OF 422-551.  
 RX MEDLINE=98043737; PubMed=9374522;  
 RA Zheng J., Chen R.H., Corbhan-Garcia S., Cahill S.M., Bar-Sagi D.,  
 RA Cowburn D.;  
 RT "The solution structure of the pleckstrin homology domain of human  
 RT SOS1. A possible structural role for the sequential association of  
 RT diffuse B cell lymphoma and pleckstrin homology domains.";  
 RL J. Biol. Chem. 272:30340-30344(1997).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 198-551.  
 RA MEDLINE=99005193; PubMed=9790532;  
 RA Solisson S.M., Mnmual A.S., Uy M., Bar-Sagi D., Kuriyan J.;  
 RT "Crystal structure of the db1 and pleckstrin homology domains from  
 RT the human Son of sevenless protein.";  
 RL Cell 95:259-268(1998).  
 CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP.  
 CC -1- SUBUNIT: INTERACTS WITH GRB2.  
 CC -1- SIMILARITY: Contains 1 db1-homology (DH) domain.  
 CC -1- SIMILARITY: Contains 1 PH domain.  
 CC -1- SIMILARITY: Contains 1 Ras-GEF domain.  
 CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: L13857; AAA35913.1; -  
 DR PDB: 1AME; 25-FEB-98.  
 DR PDB; 1DBH; 23-DEC-98.  
 DR Genew; HGNC:11187; SOS1.  
 DR MIM; 182530; -  
 DR GO: GO:0005088; F:RAS guanyl-nucleotide exchange factor activity; TAS.  
 DR GO: GO:0005100; F:Rho GTPase activator activity; TAS.  
 DR GO: GO:0005089; F:Rho guanyl-nucleotide exchange factor activity; TAS.  
 DR GO: GO:0007265; P:RAS protein signal transduction; TAS.  
 DR InterPro: IPR001331; GDS\_CDC24.  
 DR InterPro: IPR004822; Histone\_core.  
 DR InterPro: IPR001849; PH.  
 DR InterPro: IPR000651; RasGEF.  
 DR InterPro: IPR001895; RasGEF\_CDC25.  
 DR InterPro: IPR000219; RhoGEF.  
 DR Pfam; PF00169; PH; 1.  
 DR Pfam; PF00617; RasGEF; 1.  
 DR Pfam; PF00618; RasGEFN; 1.  
 DR Pfam; PF00621; RhoGEF; 1.  
 DR SMART; SM00233; PH; 1.  
 DR SMART; SM00147; RasGEF; 1.  
 DR SMART; SM00229; RasGEFN; 1.  
 DR SMART; SM00325; RhoGEF; 1.  
 DR PROSITE; PS00741; DH\_1; FALSE\_NEG.  
 DR PROSITE; PS50010; DH\_2; 1.  
 DR PROSITE; PS500720; RasGEF; 1.  
 DR PROSITE; PS50003; PH\_DOMAIN; 1.  
 DR PROSITE; PS50009; RasGEF\_CAT; 1.  
 DR PROSITE; PS50212; RasGEF\_NTER; 1.  
 DR KW Guanine-nucleotide releasing factor; 3D-structure.  
 FT DOMAIN 200 390  
 FT DOMAIN 444 548  
 FT DOMAIN 597 741  
 FT DOMAIN 780 1019  
 FT DOMAIN 1258 1261  
 FT HELIX 201 224  
 FT TURN 225 227  
 FT HELIX 228 232  
 FT TURN 234 236  
 FT HELIX 239 246  
 FT TURN 247 248  
 FT HELIX 249 268  
 FT TURN 269 269  
 FT TURN 272 273  
 FT HELIX 280 288  
 FT TURN 289 292  
 FT HELIX 293 302  
 FT TURN 303 303  
 FT TURN 305 306  
 FT HELIX 307 316  
 FT TURN 318 319  
 FT HELIX 320 327  
 FT TURN 329 330  
 FT TURN 331 337  
 FT TURN 338 338  
 FT HELIX 339 342  
 FT TURN 343 344  
 FT HELIX 345 363  
 FT TURN 367 379  
 FT TURN 380 380  
 FT HELIX 381 392  
 FT HELIX 394 402  
 FT TURN 403 403  
 FT TURN 419 419  
 FT TURN 420 426  
 FT HELIX 427 428  
 FT TURN 429 430  
 FT STRAND 437 439  
 FT HELIX 440 440  
 FT STRAND 444 452  
 FT STRAND 445 454  
 FT STRAND 459 464

```

CC STRAND 467 473
CC TURN 475 476
CC TURN 481 482
CC STRAND 488 495
CC STRAND 500 503
CC STRAND 513 516
CC TURN 519 520
CC STRAND 524 527
CC TURN 531 545
CC HELIX 546 549
CC SEQUENCE 1333 AA; 152463 MW; C6B99CCAl1ABDE45 CRC64;

Query Match 11.0%; Score 193; DB 1; Length 1333;
Best Local Similarity 27.9%; Pred. No. 2.5e-08;
Matches 60; Conservative 46; Mismatches 83; Indels 26; Gaps 9;

OY PTV-----GSAEGLDVSAR--DLAQLTDFHDSLSINHOVLHYVAGPQLRPTV 115
DB PTVEMHISRPQHIEFDLTLHPLEIARQLLLESDDLRAVQPSLVGSVWTKED-KEIN 817
OY 116 TANLEFRMRRNELQYVATELCPCVPG--PRAQLRKFKLAHLKEOKNLNSFFAVM 173
DB 818 SPNLKMIKRIHTNLTLMF--EKCIYETENLEBRVAVSRITIELQVPELNNFNCLEVV 875
OY 174 FGLSASASRLAHTERLPHVKRYLSALERLDPSSNHRVYRLALAKL---SPVIVEM 230
DB 876 SAMNSSPYRLDHFPEQIPSRQKI--LEEAHEISEDH--YKTYLAKRLSRINPCVPEF 930
OY 231 PLLKDMTFIEGHNHTLV-----ENLINFEMKRMMA 261
DB 931 GYLTLNLIKTEGNEPEVLRKRGKELINFSKRKKA 965

RESULT 12
SOS2_HUMAN STANDARD; PRT; 1332 AA.
ID SOS2_HUMAN STANDARD; PRT; 1332 AA.
AC Q07890; Q15503; (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Son of sevenless protein homolog 2 (SOS-2).
GN SOS2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=93262494; PubMed=8493579;
RA Chardin P.; Camonis J.H.; Gale N.W.; van Aelst L.; Wigler M.H.;
RA Bar-Sagi D.;
RT "Human Sos1: a guanine nucleotide exchange factor for Ras that binds
RT to GRB2."
RT Science 260:1338-1343(1993).
RN [2]
RP SEQUENCE OF 724-1296 FROM N.A.
RC TISSUE=Placenta;
RA Fath I.; Apio F.; Dutrillaux B.; Tocque B.;
RA Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP (BY
CC SIMILARITY).
CC -1- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC -1- SIMILARITY: Contains 1 PH domain.
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L13858; AAA35914.1; -
CC EMBL; L20686; AAA91852.1; -
CC HSSP; Q62245; IPMS.
CC Genew; HGNC:11188; SOS2.
CC MIM; 601247; -.
CC GO; GO:0005085; F:guanylyl-nucleotide exchange factor activity; NAS.
CC GO; GO:0007264; P:small GTPase mediated signal transduction; NAS.
CC InterPro; IPR001331; GDS_CDC24.
CC InterPro; IPR001849; PH.
CC InterPro; IPR000651; RasGEFN.
CC InterPro; IPR001895; RasGEF_CDC25.
CC InterPro; IPR000219; RhogEF.
CC Pfam; PF00169; PH; 1.
CC Pfam; PF00617; RasGEF; 1.
CC Pfam; PF00618; RasGEFN; 1.
CC Pfam; PF00621; RhogEF; 1.
CC SMART; SM00233; PH; 1.
CC SMART; SM00147; RasGEF; 1.
CC SMART; SM00229; RasGEFN; 1.
CC SMART; SM00325; RhogEF; 1.
CC PROSITE; PS00741; DH_1; FALSE_NEG.
CC PROSITE; PS50010; DH_2; 1.
CC PROSITE; PS50020; RasGEF; 1.
CC PROSITE; PS50003; PH_DOMAIN; 1.
CC PROSITE; PS50009; RasGEF_CAT; 1.
CC PROSITE; PS50212; RasGEF_NTER; 1.
CC KW Guanine-nucleotide releasing factor.
CC FT DOMAIN 198 388
CC FT 442 546
CC FT 595 739
CC FT 778 1017
CC FT 755 758
CC FT 1180 1183
CC FT DOMAIN 1203 1208
CC FT 778 778
CC FT 861 861
CC FT 948 948
CC FT 999 999
CC FT 1032 1032
CC FT 1042 1042
CC FT 1112 1114
CC FT 1159 1159
CC FT 1296 1296
CC FT 1296 1296
CC SEQUENCE 1332 AA; 153030 MW; EE4B9E52F853AC9 CRC64;

Query Match 10.7%; Score 188.5; DB 1; Length 1332;
Best Local Similarity 24.8%; Pred. No. 6e-08;
Matches 60; Conservative 53; Mismatches 102; Indels 27; Gaps 9;

OY 29 QDARGVATSGLNLRLFVNPOEVHELIPHDQLGPTVGSAGEGLDVSAR--DLAQLT 86
DB 740 QAOANGVSHNI-----TFESPPPIEMHISKP-----GQFETFDLMTLPLEIARQLT 787
OY 87 DHDMSLFNSIHQVELHYVLPQHLADYTTANLEFRMRRNELQYVATELCPCVPG-- 144
DB 788 LLESDLRYKRVQPSSELVGSVWTKED-KEINSPMLKMIKRIHTNLTLMF--EKCIYAEAEKFE 844
OY 145 PRAQLRKFKIKLAHLKEOKNLNSFFAVMFGLSNSAISRLAHTWBERLPHVKRYLSALER 204
DB 845 ERYVAVALSRITIELQVFDLNNFNGVLEIVSAVNSVYRLDHTFPALEKRRKI--LDE 901
OY 205 LIDPSWMNH-RVYRLALAKISPPVDPMPILLKDMTFIEGHNHTLV-----ENLINFEMKRM 259
DB 902 AVELSQDHFKKYLKLSINPCVPEFGIYLTNLIKTEGNNDFLRKRGKDLINFSKRK 961

OY 260 MA 261
DB 962 VA 963

RESULT 13

```

SC25\_YEAST  
ID SC25\_YEAST STANDARD: PRT: 1252 AA.  
AC P14771: Q12037: 012065;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Guanine nucleotide exchange factor SDC25.  
GN SDC25 OR YL016W/YL017W OR L1309/L1305.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae;  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M303;  
RX MEDLINE=91094833; PubMed=1986220;  
RA Damak F., Boy-Marcotte E., Le Roscouet D., Gullbaud R., Jacquet M.;  
RT "SDC25, a CDC25-like gene which contains a Ras-activating domain and  
RL Mol. Cell. Biol. 11:202-212(1991)."  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C/FT23;  
RX MEDLINE=96405918; PubMed=8810043;  
RA Miosga T., Zimmermann F.K.;  
RT "Sequence analysis of the CEN12 region of Saccharomyces cerevisiae on  
RT a 4.3.7 kb fragment of chromosome XII including an open reading frame  
RT homologous to the human cystic fibrosis transmembrane conductance  
RT regulator protein CFTR.";  
RL Yeast 12:693-708(1996)."  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE=97197984; PubMed=9046100;  
RA Purnelle B., Goffeau A.;  
RT "The sequence of 32kb on the left arm of yeast chromosome XII reveals  
RT six known genes, a new member of the seripauperins family and a new  
RT ABS transporter homologous to the human multidrug resistance  
RT protein.";  
RL Yeast 13:183-188(1997)."  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C/AB972;  
RX MEDLINE=97313267; PubMed=9169871;  
RA Johnson M., Hillier L., Riles L., Albermann K., Andre B., Ansojge W.,  
RA Benes V., Brueckner M., Deltus H., Dubois E., Duesterhoeft A.,  
RA Ertlan K.-D., Floeth M., Goffeau A., Hedling U., Heumann K.,  
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kline K., Koeltter P.,  
RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,  
RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,  
RA Portetelle D., Purnelle B., Rechner S., Rieger M., Rinke M., Rose M.,  
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,  
RA Underwood A.P., Urrutiarazu L.A., Vandenbol M., Verhasselt P.,  
RA Vierendeels F., Voet M., Voickaert G., Voss H., Wambutt E.,  
RA Wedler H., Zimmermann F.K., Zollner A., Hanl J., Hohnselt J.D.;  
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";  
RL Nature 387:87-90(1997)."  
RN [5]  
RP SEQUENCE OF 667-1252 FROM N.A.  
RC STRAIN=OL136;  
RX MEDLINE=89306677; PubMed=2545538;  
RA Boy-Marcotte E., Damak F., Camonis J., Garreau H., Jacquet M.;  
RT "The C-terminal part of a gene partially homologous to CDC 25 gene  
RT suppresses the cdc25-5 mutation in Saccharomyces cerevisiae.";  
RL Gene 77:21-30(1989)."  
RN [6]  
RP FUNCTION;  
RX MEDLINE=90260633; PubMed=2188363;  
RA Crechet J.B., Pouillet P., Mistou M.-Y., Parmeggiani A., Camonis J.,  
RA Boy-Marcotte E., Damak F., Jacquet M.;  
RT "Enhancement of the GDP-GTP exchange of Ras proteins by the carboxyl-  
RT terminal domain of SCD25.";  
RL Science 248:866-868(1990)."

RN [7]  
RP FUNCTION;  
RX MEDLINE=91156312; PubMed=2000228;  
RA Rey I., Schweighoffer F., Barlat I., Camonis J., Boy-Marcotte E.,  
RA Gullbaud R., Jacquet M., Tocque B.;  
RT "The COOH-domain of the product of the Saccharomyces cerevisiae SCD25  
RT gene elicits activation of p21-ras proteins in mammalian cells.";  
RL Oncogene 6:347-349(1991)."  
RN [8]  
RP FUNCTION;  
RX PubMed=8730097;  
RA Boy-Marcotte E., Ikonomi P., Jacquet M.;  
RT "SDC25, a dispensable Ras guanine nucleotide exchange factor of  
RT Saccharomyces cerevisiae differs from CDC25 by its regulation.";  
RL Mol. Biol. Cell 7:529-539(1996)."  
CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP.  
CC -1- MISCELLANEOUS: SUPPRESSES THE CDC25-5 MUTATION IN YEAST (RESTORES  
CC CAMP LEVEL) AND HAS SIMILAR FUNCTIONS AS CDC25.  
CC -1- MISCELLANEOUS: Strain S288C has a natural frameshift in position  
CC 91 which disrupt the gene coding for this protein and produces two  
CC ORFs.  
CC -1- SIMILARITY: Contains 1 SH3 domain.  
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.  
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M26647; AAA16565.1; -  
DR EMBL: X91488; CAA62775.1; ALT\_FRAME.  
DR EMBL: X91488; CAA62774.1; ALT\_FRAME.  
DR EMBL: X97560; CAA66173.1; ALT\_FRAME.  
DR EMBL: X97560; CAA66173.1; ALT\_FRAME.  
DR EMBL: 273121; CAA97461.1; ALT\_FRAME.  
DR EMBL: 273122; CAA97462.1; ALT\_FRAME.  
DR SGD: S0003939; SDC25.  
DR InterPro: IPR000651; RASGEFN.  
DR InterPro: IPR001895; RASGRF\_CDC25.  
DR InterPro: IPR001452; SH3.  
DR Pfam: PF006617; RASGEF.1.  
DR Pfam: PF006618; RASGEF.1.  
DR Pfam: PF00018; SH3.1.  
DR SMART: SM00147; RASGEF.1.  
DR SMART: SM00229; RASGEF.1.  
DR SMART: SM00326; SH3.1.  
DR PROSITE: PS00720; RASGEF.1.  
DR PROSITE: PS50009; RASGEF\_CAT.1.  
DR PROSITE: PS50212; RASGEF\_NTFR.1.  
DR PROSITE: PS50002; SH3.1.  
KW Guanine-nucleotide releasing factor; Cell cycle; Cell division;  
KW SH3 domain.  
FT DOMAIN 26 97  
FT DOMAIN 782 914  
FT DOMAIN 952 1199  
FT DOMAIN 74 78  
FT DOMAIN 433 436  
FT VARIANT 583 589  
FT CONFLICT 78 78  
FT CONFLICT 304 304  
FT CONFLICT 377 379  
FT CONFLICT 385 386  
FT CONFLICT 417 418  
FT CONFLICT 553 553  
FT CONFLICT 557 557  
FT CONFLICT 782 784  
FT CONFLICT 809 809  
FT CONFLICT 875 875  
FT CONFLICT 915 915  
SH3.  
N-TERMINAL RAS-GEF.  
RAS-GEF.  
POLY-ASN.  
POLY-ARG.  
DVVVKFI -> V (IN STRAIN OL136).  
N -> NN (IN REF. 1).  
L -> S (IN REF. 1).  
KRL -> IVF (IN REF. 1).  
HD -> MT (IN REF. 1).  
SC -> FI (IN REF. 1).  
D -> G (IN REF. 1).  
E -> K (IN REF. 1).  
SNM -> PIV (IN REF. 1 AND 5).  
D -> N (IN REF. 1 AND 5).  
L -> P (IN REF. 1 AND 5).  
H -> L (IN REF. 1 AND 5)."





```

Db 1310 ANQLVLEHDLVLRITMFECLDRAMCTKYCNMGSPNITKFIANANTLNFVSHITVKA 1369
QY 142 VPEPRAQLRKFKLAHLKEOKNLNSFFAVMGLNSAISRLAHTMERLPHVKRLYSA 201
Db 1370 DKTRSKLTQYFTVAQHCKELNLFSSMRIVSALYSSPIRLKTKWDVSTSKDLKN 1429
QY 202 LERLLDPSNNHRYRLALAKLSP-PIVPMPLLLKDMTEIHGNNHTLVE--NLINEKM 257
Db 1430 LNNLMDSKRNFFVYRELRSVTDVACVPEFGYVLSLDTFTVCNPDLNSTNINPSKR 1489
QY 258 RKMA 261
Db 1490 TKIA 1493

RESULT 15
SOS_DROME STANDARD; PRT; 1595 AA.
ID SOS_DROME
AC P26675;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Son of sevenless protein.
GN SOS.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID:7227;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-Oregon-R;
RX MEDLINE-92141820; Pubmed-1736363;
RA Bonfini L., Karlovich C.A., Dasgupta C., Banerjee U.;
RT "The Son of sevenless gene product: a putative activator of Ras.";
RL Science 253:603-606(1992).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE-92034991; Pubmed-1934068;
RA Simon M.A., Bottell D.D.L., Dodson G.S., Laverly T.R., Rubin G.M.;
RT "Ras1 and a putative guanine nucleotide exchange factor perform
RT crucial steps in signaling by the sevenless protein tyrosine
RT kinase.";
RL Cell 67:701-716(1991).
CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP. SOS IS
CC IMPLICATED IN NEURONAL DEVELOPMENT.
CC -1- SUBUNIT: MAY FORM A COMPLEX WITH SEVENLESS AND DRK.
CC -1- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC -1- SIMILARITY: Contains 1 Ras-GEF domain.
CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M83931; AAB04680.1; -
CC EMBL: M77501; AAA28904.1; -
CC PDB: 1AZE; 18-MAY-99.
CC FLYBase: FBgn0001965; Soc.
CC InterPro: IPR001331; GDS_CDC24.
CC InterPro: IPR004822; Histone_core.
CC InterPro: IPR002119; Histone_H2A.
CC InterPro: IPR001849; PH.
CC InterPro: IPR000651; RasGEFN.
CC InterPro: IPR001895; RasGRF_CDC25.
CC InterPro: IPR000219; RhoGEF.
CC Pfam: PF00169; PH; 1.
CC Pfam: PF00617; RasGEF; 1.

```

```

DR Pfam; PF00618; RasGEFN; 1.
DR Pfam; PF00621; RhoGEF; 1.
DR PRINTS; PR00620; HISTONEH2A.
DR SMART; SM00233; PH; 1.
DR SMART; SM00147; RasGEFN; 1.
DR SMART; SM00229; RasGEFN; 1.
DR SMART; SM00325; RhoGEF; 1.
DR PROSITE; PS00741; DH_1; FALSE_NEG.
DR PROSITE; PS50010; DH_2; 1.
DR PROSITE; PS00720; RasGEF; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50009; RasGEF_CAT; 1.
DR PROSITE; PS50212; RasGEF_NTER; 1.
KW Guanine-nucleotide releasing factor; Neurogenesis; 3D-structure.
FT DOMAIN 636 791 N-TERMINAL RAS-GEF.
FT DOMAIN 828 1065 RAS-GEF.
FT DOMAIN 15 51 GLY-RICH.
FT DOMAIN 247 433 DH.
FT DOMAIN 479 587 PH.
FT DOMAIN 1511 1516 GLN-RICH.
FT DOMAIN 1525 1541 HIS-RICH.
FT CONFLICT 232 243 TSCPVPCHPERS -> HILSPPLSLPAQR
FT CONFLICT (IN REF. 2).
FT CONFLICT V -> P (IN REF. 2).
SQ SEQUENCE 1595 AA; 177837 MW; 33AE31F0767A219F CRC64;

Query Match 10.1%; Score 178; DB 1; Length 1595;
Best Local Similarity 25.3%; Pred. No. 6.1e-07;
Matches 55; Conservative 47; Mismatches 95; Indels 20; Gaps 6;

QY 50 POEVELIHPHDOLGPTVGSABGLDLYSANDLAGOLTDHDSLFNSIHQVELIHVYVGPQ 109
Db 810 PIHHLISVPRDETLTLIHPLE-----LARQLLLEFEWYKNVKKPSLVG---SPW 857
QY 110 HLBD--VTANLERFRMRNFELQYVATELCICLPVCEPRAQLRKFKLAHLKEOKNLN 167
Db 858 TKRDKYKSPNLKIKKHTTNTVTRWIEKSTEAENVEERLAINQRALEVMMVMELENNFN 917
QY 168 SFAVVMGLNSAISRLAHTMERLPHVKRLYSALERTLDPSMNH-RVYRLALAKLSPV 226
Db 918 GLIStYAAMGTASVYRLRWTFQGLPERYKRF---LECCRSLSDHLKRYERLRSINPFC 974
QY 227 IPPMPLLLKDMTEIHGNNHTLVEN--LINEKRMMA 261
Db 975 VPEFGRYLTNIHLDEGNPDLANTLINSKRKRYA 1011

```

Search completed: October 7, 2003, 07:13:42  
Job time : 27 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 07:13:18 : Search time 402 Seconds  
(without alignments)  
765.056 Million cell updates/sec

Title: US-09-856-679-2  
Perfect score: 1759  
Sequence: 1 MAALAQEDGWTKGVLVKVN.....QLKVIDNQRELSRLSRELEP 338

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*  
1: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US09\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US09\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US09\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US09\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US09\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US09\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US09\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US09\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
32: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*  
33: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1759	100.0	338	23	US-09-856-679-2 Sequence 2, Appl1

2	1759	100.0	881	32	US-60-452-680-21605
3	1759	100.0	881	32	US-60-455-444-7029
4	1759	100.0	881	32	US-60-461-762-180
5	1759	100.0	881	32	US-60-465-241-7029
6	1759	100.0	881	32	US-60-470-166-1217
7	1759	100.0	923	1	PCT-US02-36151-29
8	1759	100.0	923	32	US-60-452-680-21606
9	1759	100.0	923	32	US-60-455-444-7030
10	1759	100.0	923	32	US-60-461-762-181
11	1759	100.0	923	32	US-60-465-241-7030
12	1759	100.0	923	32	US-60-470-166-1218
13	1755	99.8	881	22	US-09-791-537-29766
14	1748	99.4	881	18	US-09-422-999-12
15	1748	99.4	881	22	US-09-791-537-64901
16	1664	94.6	884	18	US-09-422-999-10
17	1664	94.6	884	22	US-09-791-537-64922
18	1644.5	93.5	897	21	US-09-724-676-84651
19	1644.5	93.5	897	21	US-09-724-676-84651
20	1643	93.4	899	32	US-60-245-228-224
21	1563	88.9	592	30	US-10-408-765A-1711
22	1563	88.9	592	32	US-60-389-987-1711
23	1563	88.9	592	32	US-60-412-418-1711
24	1408	80.0	270	24	US-09-911-826A-8
25	1013	57.6	235	22	US-09-760-464A-1834
26	1013	57.6	235	22	US-09-760-485-726
27	1013	57.6	235	28	US-10-206-664-1834
28	1013	57.6	235	28	US-10-216-436-726
29	863.5	49.1	552	32	US-60-245-228-298
30	863.5	49.1	552	32	US-60-258-272-124
31	863.5	49.1	849	18	US-09-422-999-18
32	863.5	49.1	1011	22	US-09-791-537-108781
33	863.5	49.1	1011	28	US-10-219-051B-8943
34	863.5	49.1	1011	28	US-10-219-051B-8943
35	855.5	48.6	993	22	US-09-791-537-118055
36	832.5	47.3	436	22	US-09-791-537-118055
37	832.5	47.3	436	28	US-10-219-051B-8943
38	832.5	47.3	436	28	US-10-219-051B-8943
39	786.5	44.7	291	18	US-09-422-999-16
40	696	39.6	522	32	US-60-212-664-431
41	696	39.6	957	20	US-09-614-150-6456
42	696	39.6	957	22	US-09-791-537-29071
43	696	39.6	957	32	US-60-167-217-5571
44	696	39.6	957	32	US-60-173-464-5273
45	696	39.6	957	32	US-60-191-637-6477

#### ALIGNMENTS

RESULT 1  
US-09-856-679-2  
Sequence 2 Application US/09856679  
GENERAL INFORMATION:  
APPLICANT: INCYTE PHARMACEUTICALS, INC.  
APPLICANT: HILLMAN, Jennifer L.  
APPLICANT: TANG, Y. Tom  
APPLICANT: BANDMAN, Olga  
APPLICANT: LAL, Preeti  
APPLICANT: YUE, Henry  
APPLICANT: LU, Dying Aina M.  
APPLICANT: BAUGHN, Mariah R.  
APPLICANT: YANG, Junming  
APPLICANT: AZIMZAI, Valda  
TITLE OF INVENTION: GTPASE ASSOCIATED PROTEINS  
FILE REFERENCE: PF-0629 PCT  
CURRENT APPLICATION NUMBER: US/09/856,679  
PRIOR FILING DATE: 2002-01-23  
PRIOR APPLICATION NUMBER: 60/109,592  
PRIOR FILING DATE: 1998-11-23  
PRIOR APPLICATION NUMBER: 60/118,610  
PRIOR FILING DATE: 1999-02-04  
PRIOR APPLICATION NUMBER: 60/127,990  
PRIOR FILING DATE: 1999-04-06

Sequence 21605, A  
Sequence 7029, Ap  
Sequence 480, Ap  
Sequence 7029, Ap  
Sequence 1217, Ap  
Sequence 29, Appl  
Sequence 21606, A  
Sequence 7030, Ap  
Sequence 481, Ap  
Sequence 7030, Ap  
Sequence 1218, Ap  
Sequence 29766, A  
Sequence 12, Appl  
Sequence 10, Appl  
Sequence 64922, A  
Sequence 84651, A  
Sequence 84651, A  
Sequence 224, Ap  
Sequence 1711, Ap  
Sequence 1711, Ap  
Sequence 8, Appl  
Sequence 1834, Ap  
Sequence 726, Ap  
Sequence 1834, Ap  
Sequence 726, Ap  
Sequence 299, Ap  
Sequence 124, Ap  
Sequence 18781, A  
Sequence 8943, Ap  
Sequence 8947, Ap  
Sequence 118035, A  
Sequence 18035, A  
Sequence 8945, Ap  
Sequence 431, Ap  
Sequence 431, Ap  
Sequence 6456, Ap  
Sequence 29071, A  
Sequence 6571, Ap  
Sequence 5273, Ap  
Sequence 6477, Ap

```
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 338
; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: 1259937CD1
US-09-856-679-2
```

```
Query Match      100.0%: Score 1759; DB 23; Length 338;
Best Local Similarity 100.0%: Pred. No. 4.3e-173;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MAALAQEDGWTGQVLYVNSAGDAIGLQDPARGVATSLGINERLFVVPQEVHELIPHP 60
DB 1 MAALAQEDGWTGQVLYVNSAGDAIGLQDPARGVATSLGINERLFVVPQEVHELIPHP 60
OY 61 DQLGPTVSAGGLDLYSAKDLAAGQLTDHWSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
DB 61 DQLGPTVSAGGLDLYSAKDLAAGQLTDHWSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
OY 121 RFRMRFNELQYVATLCLCPVPGRAQLRKFTKLAHLKEOKNLNSFFAVMGSLNSA 180
DB 121 RFRMRFNELQYVATLCLCPVPGRAQLRKFTKLAHLKEOKNLNSFFAVMGSLNSA 180
OY 181 ISRLAHTWERLPHKVRKLYSALERLDPSSMHRVRYRLALAKSPVPIPFMPLLLKDMFTI 240
DB 181 ISRLAHTWERLPHKVRKLYSALERLDPSSMHRVRYRLALAKSPVPIPFMPLLLKDMFTI 240
OY 241 HEGNHTLVENLINEKMMARARMLHCHSHNPVPLSPRSRVSHLHEDSOVARISTC 300
DB 241 HEGNHTLVENLINEKMMARARMLHCHSHNPVPLSPRSRVSHLHEDSOVARISTC 300
OY 301 SEQSLSTRSPASTWAVYQOLKVIDNORELSRSLRELP 338
DB 301 SEQSLSTRSPASTWAVYQOLKVIDNORELSRSLRELP 338
```

```
RESULT 2
US-60-452-680-21605
; Sequence 21605, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GROPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21605
; LENGTH: 881
; TYPE: prt
; ORGANISM: Homo sapiens
US-60-452-680-21605
```

```
Query Match      100.0%: Score 1759; DB 32; Length 881;
Best Local Similarity 100.0%: Pred. No. 1.9e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MAALAQEDGWTGQVLYVNSAGDAIGLQDPARGVATSLGINERLFVVPQEVHELIPHP 60
DB 544 MAALAQEDGWTGQVLYVNSAGDAIGLQDPARGVATSLGINERLFVVPQEVHELIPHP 603
OY 61 DQLGPTVSAGGLDLYSAKDLAAGQLTDHWSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
DB 604 DQLGPTVSAGGLDLYSAKDLAAGQLTDHWSLFNSIHQVELIHVYLGPOHLRDVTTANLE 663
OY 121 RFRMRFNELQYVATLCLCPVPGRAQLRKFTKLAHLKEOKNLNSFFAVMGSLNSA 180
DB 121 RFRMRFNELQYVATLCLCPVPGRAQLRKFTKLAHLKEOKNLNSFFAVMGSLNSA 180
```

```
DB 664 RFRMRFNELQYVATLCLCPVPGRAQLRKFTKLAHLKEOKNLNSFFAVMGSLNSA 723
OY 181 ISRLAHTWERLPHKVRKLYSALERLDPSSMHRVRYRLALAKSPVPIPFMPLLLKDMFTI 240
DB 724 ISRLAHTWERLPHKVRKLYSALERLDPSSMHRVRYRLALAKSPVPIPFMPLLLKDMFTI 783
OY 241 HEGNHTLVENLINEKMMARARMLHCHSHNPVPLSPRSRVSHLHEDSOVARISTC 300
DB 784 HEGNHTLVENLINEKMMARARMLHCHSHNPVPLSPRSRVSHLHEDSOVARISTC 843
OY 301 SEQSLSTRSPASTWAVYQOLKVIDNORELSRSLRELP 338
DB 844 SEQSLSTRSPASTWAVYQOLKVIDNORELSRSLRELP 881
```

```
RESULT 3
US-60-455-444-7029
; Sequence 7029, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001455
; CURRENT APPLICATION NUMBER: US/60/455,444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7029
; LENGTH: 881
; TYPE: prt
; ORGANISM: Homo sapiens
US-60-455-444-7029
```

```
Query Match      100.0%: Score 1759; DB 32; Length 881;
Best Local Similarity 100.0%: Pred. No. 1.9e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MAALAQEDGWTGQVLYVNSAGDAIGLQDPARGVATSLGINERLFVVPQEVHELIPHP 60
DB 544 MAALAQEDGWTGQVLYVNSAGDAIGLQDPARGVATSLGINERLFVVPQEVHELIPHP 603
OY 61 DQLGPTVSAGGLDLYSAKDLAAGQLTDHWSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
DB 604 DQLGPTVSAGGLDLYSAKDLAAGQLTDHWSLFNSIHQVELIHVYLGPOHLRDVTTANLE 663
OY 121 RFRMRFNELQYVATLCLCPVPGRAQLRKFTKLAHLKEOKNLNSFFAVMGSLNSA 180
DB 664 RFRMRFNELQYVATLCLCPVPGRAQLRKFTKLAHLKEOKNLNSFFAVMGSLNSA 723
OY 181 ISRLAHTWERLPHKVRKLYSALERLDPSSMHRVRYRLALAKSPVPIPFMPLLLKDMFTI 240
DB 724 ISRLAHTWERLPHKVRKLYSALERLDPSSMHRVRYRLALAKSPVPIPFMPLLLKDMFTI 783
OY 241 HEGNHTLVENLINEKMMARARMLHCHSHNPVPLSPRSRVSHLHEDSOVARISTC 300
DB 784 HEGNHTLVENLINEKMMARARMLHCHSHNPVPLSPRSRVSHLHEDSOVARISTC 843
OY 301 SEQSLSTRSPASTWAVYQOLKVIDNORELSRSLRELP 338
DB 844 SEQSLSTRSPASTWAVYQOLKVIDNORELSRSLRELP 881
```

```
RESULT 4
US-60-461-762-480
; Sequence 480, Application US/60461762
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GROPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001459
; CURRENT APPLICATION NUMBER: US/60/461,762
```

```

; CURRENT FILING DATE: 2003-04-11
; NUMBER OF SEQ ID NOS: 30297
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 480
; LENGTH: 881
; TYPE: PRF
; ORGANISM: Homo sapiens
US-60-461-762-480

```

```

Query Match      100.0%; Score 1759; DB 32; Length 881;
Best Local Similarity 100.0%; Pred. No. 1.9e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MAAAGDEGWTGQGVYLVKNSAGDAGLQDPDARGVATSLGNERLFFVNPQEVNELIPHP 60
DB 544 MAAAGDEGWTGQGVYLVKNSAGDAGLQDPDARGVATSLGNERLFFVNPQEVNELIPHP 603
QY 61 DOLGPTVGSAGEGLDVSADKDLAQULTDHDMSLFNSIHQVELIHYVLGPQHLRDVTYANLE 120
DB 604 DOLGPTVGSAGEGLDVSADKDLAQULTDHDMSLFNSIHQVELIHYVLGPQHLRDVTYANLE 663
QY 121 RFRARRNELOYWATVETLCPCVPGPRAQLLRKFTIKLAHLKEQKLNSEFFAVMEGLSNA 180
DB 664 RFRARRNELOYWATVETLCPCVPGPRAQLLRKFTIKLAHLKEQKLNSEFFAVMEGLSNA 723
QY 181 ISRLATWTERLPHKVKLYSALERLDPSPMNRHVRYLALAKSPVPIPFMPPLLLKDMFTI 240
DB 724 ISRLATWTERLPHKVKLYSALERLDPSPMNRHVRYLALAKSPVPIPFMPPLLLKDMFTI 783
QY 241 HEGNHTLVENLINFEMRMMAARAARMLHHCRRSHNPVPLSPLSRVSHLHEDSOVARISTC 300
DB 784 HEGNHTLVENLINFEMRMMAARAARMLHHCRRSHNPVPLSPLSRVSHLHEDSOVARISTC 843
QY 301 SEOSLSTRSPASTWATVVOOLKVIDNORELSRSLRELEP 338
DB 844 SEOSLSTRSPASTWATVVOOLKVIDNORELSRSLRELEP 881

```

# RESULT 5 US-60-465-241-7029

```

; Sequence 7029, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/465, 241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7029
; LENGTH: 881
; TYPE: PRF
; ORGANISM: Homo sapiens
US-60-465-241-7029

```

```

Query Match      100.0%; Score 1759; DB 32; Length 881;
Best Local Similarity 100.0%; Pred. No. 1.9e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MAAAGDEGWTGQGVYLVKNSAGDAGLQDPDARGVATSLGNERLFFVNPQEVNELIPHP 60
DB 544 MAAAGDEGWTGQGVYLVKNSAGDAGLQDPDARGVATSLGNERLFFVNPQEVNELIPHP 603
QY 61 DOLGPTVGSAGEGLDVSADKDLAQULTDHDMSLFNSIHQVELIHYVLGPQHLRDVTYANLE 120
DB 604 DOLGPTVGSAGEGLDVSADKDLAQULTDHDMSLFNSIHQVELIHYVLGPQHLRDVTYANLE 663
QY 121 RFRARRNELOYWATVETLCPCVPGPRAQLLRKFTIKLAHLKEQKLNSEFFAVMEGLSNA 180
DB 664 RFRARRNELOYWATVETLCPCVPGPRAQLLRKFTIKLAHLKEQKLNSEFFAVMEGLSNA 723

```

```

QY 181 ISRLATWTERLPHKVKLYSALERLDPSPMNRHVRYLALAKSPVPIPFMPPLLLKDMFTI 240
DB 724 ISRLATWTERLPHKVKLYSALERLDPSPMNRHVRYLALAKSPVPIPFMPPLLLKDMFTI 783
QY 241 HEGNHTLVENLINFEMRMMAARAARMLHHCRRSHNPVPLSPLSRVSHLHEDSOVARISTC 300
DB 784 HEGNHTLVENLINFEMRMMAARAARMLHHCRRSHNPVPLSPLSRVSHLHEDSOVARISTC 843
QY 301 SEOSLSTRSPASTWATVVOOLKVIDNORELSRSLRELEP 338
DB 844 SEOSLSTRSPASTWATVVOOLKVIDNORELSRSLRELEP 881

```

# RESULT 6 US-60-470-166-1217

```

; Sequence 1217, Application US/60470166
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/470, 166
; CURRENT FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 68617
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1217
; LENGTH: 881
; TYPE: PRF
; ORGANISM: Homo sapiens
US-60-470-166-1217

```

```

Query Match      100.0%; Score 1759; DB 32; Length 881;
Best Local Similarity 100.0%; Pred. No. 1.9e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MAAAGDEGWTGQGVYLVKNSAGDAGLQDPDARGVATSLGNERLFFVNPQEVNELIPHP 60
DB 544 MAAAGDEGWTGQGVYLVKNSAGDAGLQDPDARGVATSLGNERLFFVNPQEVNELIPHP 603
QY 61 DOLGPTVGSAGEGLDVSADKDLAQULTDHDMSLFNSIHQVELIHYVLGPQHLRDVTYANLE 120
DB 604 DOLGPTVGSAGEGLDVSADKDLAQULTDHDMSLFNSIHQVELIHYVLGPQHLRDVTYANLE 663
QY 121 RFRARRNELOYWATVETLCPCVPGPRAQLLRKFTIKLAHLKEQKLNSEFFAVMEGLSNA 180
DB 664 RFRARRNELOYWATVETLCPCVPGPRAQLLRKFTIKLAHLKEQKLNSEFFAVMEGLSNA 723
QY 181 ISRLATWTERLPHKVKLYSALERLDPSPMNRHVRYLALAKSPVPIPFMPPLLLKDMFTI 240
DB 724 ISRLATWTERLPHKVKLYSALERLDPSPMNRHVRYLALAKSPVPIPFMPPLLLKDMFTI 783
QY 241 HEGNHTLVENLINFEMRMMAARAARMLHHCRRSHNPVPLSPLSRVSHLHEDSOVARISTC 300
DB 784 HEGNHTLVENLINFEMRMMAARAARMLHHCRRSHNPVPLSPLSRVSHLHEDSOVARISTC 843
QY 301 SEOSLSTRSPASTWATVVOOLKVIDNORELSRSLRELEP 338
DB 844 SEOSLSTRSPASTWATVVOOLKVIDNORELSRSLRELEP 881

```

# RESULT 7 PCT-US02-36151-29

```

; Sequence 29, Application PC/TUS0236151
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: KABE, Amy E.; SMARNKAR, Anita
; APPLICANT: GORVAD, Ann E.; HAFALIA, April J.A.
; APPLICANT: DUGGAN, Brendan M.; WARREN, Bridget A.
; APPLICANT: EMERLING, Brooke E.; ISON, Craig H.
; APPLICANT: NGUYEN, Daniel B.; LINDQUIST, Erika A.
; APPLICANT: LEE, Ernestine A.; YUE, Henry
; APPLICANT: YUE, Huibin; FORSYTHE, Ian J.

```

```

: APPLICANT: RAMKUMAR, Jayalaxmi; GRIFFIN, Jennifer A.
: APPLICANT: Li, Joanna X.; MARQUIS, Joseph P.
: APPLICANT: GIETZEN, Kimberly J.; BAUGHN, Mariah R.
: APPLICANT: BOROMSKY, Mark L.; YAO, Monique G.
: APPLICANT: CHAWLA, Narinder K.; LEHR-MASON, Patricia M.
: APPLICANT: LAL, Preeti G.; GURURAJAN, Rajagopal
: APPLICANT: KHARE, Reena; BARRA, Saijeev
: APPLICANT: BECHA, Shanya D.; LEE, Soo Yeun
: APPLICANT: TRAN, Uyen K.; ELLIOTT, Vicki S.
: APPLICANT: SPRAGUE, William W.; TANG, Y. Tom
: APPLICANT: ZEBARJADIAN, Yeganeh; JIANG, Xin
: APPLICANT: JACKSON, Alan A.; BHATTI, Umesh G.
: APPLICANT: BURRILL, John D.; LEE, Sally
: APPLICANT: BLAKE, Julie J.; HO, Anne
: TITLE OF INVENTION: INTRACELLULAR SIGNALING MOLECULES
: FILE REFERENCE: PF-1275 PCT
: CURRENT APPLICATION NUMBER: PCT/US02/36151
: PRIOR FILING DATE: 2002-11-07
: PRIOR APPLICATION NUMBER: US 60/344,472
: PRIOR FILING DATE: 2001-11-09
: PRIOR APPLICATION NUMBER: US 60/334,558
: PRIOR FILING DATE: 2001-11-30
: PRIOR APPLICATION NUMBER: US 60/340,296
: PRIOR FILING DATE: 2001-12-14
: PRIOR APPLICATION NUMBER: US 60/343,557
: PRIOR FILING DATE: 2001-12-21
: PRIOR APPLICATION NUMBER: US 60/350,420
: PRIOR FILING DATE: 2002-01-18
: PRIOR APPLICATION NUMBER: US 60/351,927
: PRIOR FILING DATE: 2002-01-25
: NUMBER OF SEQ ID NOS: 104
: SOFTWARE: PERL Program
: SEQ ID NO 29
: LENGTH: 923
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID NO: 1004901CD1
: PCT-US02-36151-29

Query Match      100.0%; Score 1759; DB 1; Length 923;
Best Local Similarity 100.0%; Pred. No. 2e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

: 1 MAALAQEDGWTGKGVLYKVNAGDAIGLQDPARGVATSLGNERLFVYVNPQEVHELIPHP 60
: 586 MAALAQEDGWTGKGVLYKVNAGDAIGLQDPARGVATSLGNERLFVYVNPQEVHELIPHP 645

: 61 DOLGPTVGSARGDLVSAGKDIAGOLTDHDSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
: 646 DOLGPTVGSARGDLVSAGKDIAGOLTDHDSLFNSIHQVELIHVYLGPOHLRDVTTANLE 705

: 121 RMRFRNFELQYVWVATELCVPGPRROLLRKFTIKLAHLEKQKNLSFFAVMFGLSNSA 180
: 706 RMRFRNFELQYVWVATELCVPGPRROLLRKFTIKLAHLEKQKNLSFFAVMFGLSNSA 765

: 181 ISRLAHTWERLPHKVRKLYSALERLDPNMHRVRYRLALAKLSPVPIPFMPLDKDMTFI 240
: 766 ISRLAHTWERLPHKVRKLYSALERLDPNMHRVRYRLALAKLSPVPIPFMPLDKDMTFI 825

: 241 HEGNHTLVENLINFEKMKMARAARMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 300
: 826 HEGNHTLVENLINFEKMKMARAARMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 885

: 301 SEOSLSTRSPASTAWYVOOLKVINDQRELSTRLEP 338
: 886 SEOSLSTRSPASTAWYVOOLKVINDQRELSTRLEP 923

RESULT 8
US-60-452-680-21606
```

```

: Sequence 21606, Application US/60452680
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: GRUPE, Andrew
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CU001450
: CURRENT APPLICATION NUMBER: US/60/452,680
: CURRENT FILING DATE: 2003-03-07
: NUMBER OF SEQ ID NOS: 116213
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 21606
: LENGTH: 923
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-60-452-680-21606
```

```

Query Match      100.0%; Score 1759; DB 32; Length 923;
Best Local Similarity 100.0%; Pred. No. 2e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

: 1 MAALAQEDGWTGKGVLYKVNAGDAIGLQDPARGVATSLGNERLFVYVNPQEVHELIPHP 60
: 586 MAALAQEDGWTGKGVLYKVNAGDAIGLQDPARGVATSLGNERLFVYVNPQEVHELIPHP 645

: 61 DOLGPTVGSARGDLVSAGKDIAGOLTDHDSLFNSIHQVELIHVYLGPOHLRDVTTANLE 120
: 646 DOLGPTVGSARGDLVSAGKDIAGOLTDHDSLFNSIHQVELIHVYLGPOHLRDVTTANLE 705

: 121 RMRFRNFELQYVWVATELCVPGPRROLLRKFTIKLAHLEKQKNLSFFAVMFGLSNSA 180
: 706 RMRFRNFELQYVWVATELCVPGPRROLLRKFTIKLAHLEKQKNLSFFAVMFGLSNSA 765

: 181 ISRLAHTWERLPHKVRKLYSALERLDPNMHRVRYRLALAKLSPVPIPFMPLDKDMTFI 240
: 766 ISRLAHTWERLPHKVRKLYSALERLDPNMHRVRYRLALAKLSPVPIPFMPLDKDMTFI 825

: 241 HEGNHTLVENLINFEKMKMARAARMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 300
: 826 HEGNHTLVENLINFEKMKMARAARMLHCRSHNPVPLSPRSRVSHLHEDSOVARISTC 885

: 301 SEOSLSTRSPASTAWYVOOLKVINDQRELSTRLEP 338
: 886 SEOSLSTRSPASTAWYVOOLKVINDQRELSTRLEP 923

RESULT 9
US-60-455-444-7030
: Sequence 7030, Application US/60455444
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: BEGOVICH, Ann
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CU001455
: CURRENT APPLICATION NUMBER: US/60/455,444
: CURRENT FILING DATE: 2003-03-18
: NUMBER OF SEQ ID NOS: 50986
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 7030
: LENGTH: 923
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-60-455-444-7030

Query Match      100.0%; Score 1759; DB 32; Length 923;
Best Local Similarity 100.0%; Pred. No. 2e-172;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

: 1 MAALAQEDGWTGKGVLYKVNAGDAIGLQDPARGVATSLGNERLFVYVNPQEVHELIPHP 60
: 586 MAALAQEDGWTGKGVLYKVNAGDAIGLQDPARGVATSLGNERLFVYVNPQEVHELIPHP 645
```



```
Db      646 DOLGPTVGSAGELDVSAKDLAAGLTLDHDMSLFNSIHVELIHYVLGPOHLRDVTANLE 705
OY      121 REMRRFNELOVWVATELCLCPVGPRAQLLRKFILAAHLKEOKNLNSFFAVMGLSNSA 180
Db      706 REMRRFNELOVWVATELCLCPVGPRAQLLRKFILAAHLKEOKNLNSFFAVMGLSNSA 765
OY      181 ISRLAHTWERLPHKRYKLYSALERLLDPSMNRHYRYRLALAKISPPVYIPFMPILLKDMFTI 240
Db      766 ISRLAHTWERLPHKRYKLYSALERLLDPSMNRHYRYRLALAKISPPVYIPFMPILLKDMFTI 825
OY      241 HEGNHTLVENLINFEXKMMARAARMLHHCSSHNPVPLSPLSRSVSHLHEDSOVARISTC 300
Db      826 HEGNHTLVENLINFEXKMMARAARMLHHCSSHNPVPLSPLSRSVSHLHEDSOVARISTC 885
OY      301 SEOSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 338
Db      886 SEOSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 923

RESULT 13
US-09-791-537-29766
; Sequence 29766, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Derek
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 29766
; LENGTH: 881
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-29766

Query Match      99.8%; Score 1755; DB 22; Length 881;
Best Local Similarity 99.7%; Pred. No. 4.9e-172;
Matches 337; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 MAALAOEDGWTGGOVLVYVNSAGDAIGLOPDARGVATSLGNERLFVYVNPQEVHELIPHP 60
Db      544 MAALAOEDGWTGGOVLVYVNSAGDAIGLOPDARGVATSLGNERLFVYVNPQEVHELIPHP 603
OY      61 DOLGPTVGSAGELDVSAKDLAAGLTLDHDMSLFNSIHVELIHYVLGPOHLRDVTANLE 120
Db      604 DOLGPTVGSAGELDVSAKDLAAGLTLDHDMSLFNSIHVELIHYVLGPOHLRDVTANLE 663
OY      121 REMRRFNELOVWVATELCLCPVGPRAQLLRKFILAAHLKEOKNLNSFFAVMGLSNSA 180
Db      664 REMRRFNELOVWVATELCLCPVGPRAQLLRKFILAAHLKEOKNLNSFFAVMGLSNSA 723
OY      181 ISRLAHTWERLPHKRYKLYSALERLLDPSMNRHYRYRLALAKISPPVYIPFMPILLKDMFTI 240
Db      724 ISRLAHTWERLPHKRYKLYSALERLLDPSMNRHYRYRLALAKISPPVYIPFMPILLKDMFTI 783
OY      241 HEGNHTLVENLINFEXKMMARAARMLHHCSSHNPVPLSPLSRSVSHLHEDSOVARISTC 300
Db      784 HEGNHTLVENLINFEXKMMARAARMLHHCSSHNPVPLSPLSRSVSHLHEDSOVARISTC 843
OY      301 SEOSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 338
Db      844 SEOSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 881

RESULT 14
US-09-422-999-12
; Sequence 12, Application US/09422999
; GENERAL INFORMATION:
; APPLICANT: Kawasaki, Hiroaki
```

```
; APPLICANT: Graybiel, Ann
; APPLICANT: Housman, David
; TITLE OF INVENTION: Genes Integrating Signal Transduction Pathways
; FILE REFERENCE: MIT-103
; CURRENT APPLICATION NUMBER: US/09/422,999
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: US 60/105,507
; EARLIER FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: US 60/108,685
; EARLIER FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 12
; LENGTH: 881
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-422-999-12

Query Match      99.4%; Score 1748; DB 18; Length 881;
Best Local Similarity 99.1%; Pred. No. 2.6e-171;
Matches 335; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY      1 MAALAOEDGWTGGOVLVYVNSAGDAIGLOPDARGVATSLGNERLFVYVNPQEVHELIPHP 60
Db      544 MAALAOEDGWTGGOVLVYVNSAGDAIGLOPDARGVATSLGNERLFVYVNPQEVHELIPHP 603
OY      61 DOLGPTVGSAGELDVSAKDLAAGLTLDHDMSLFNSIHVELIHYVLGPOHLRDVTANLE 120
Db      604 DOLGPTVGSAGELDVSAKDLAAGLTLDHDMSLFNSIHVELIHYVLGPOHLRDVTANLE 663
OY      121 REMRRFNELOVWVATELCLCPVGPRAQLLRKFILAAHLKEOKNLNSFFAVMGLSNSA 180
Db      664 REMRRFNELOVWVATELCLCPVGPRAQLLRKFILAAHLKEOKNLNSFFAVMGLSNSP 723
OY      181 ISRLAHTWERLPHKRYKLYSALERLLDPSMNRHYRYRLALAKISPPVYIPFMPILLKDMFTI 240
Db      724 ISRLAHTWERLPHKRYKLYSALERLLDPSMNRHYRYRLALAKISPPVYIPFMPILLKDMFTI 783
OY      241 HEGNHTLVENLINFEXKMMARAARMLHHCSSHNPVPLSPLSRSVSHLHEDSOVARISTC 300
Db      784 HEGNHTLVENLINFEXKMMARAARMLHHCSSHNPVPLSPLSRSVSHLHEDSOVARISTC 843
OY      301 SEOSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 338
Db      844 SEOSLSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 881

RESULT 15
US-09-791-537-64901
; Sequence 64901, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Derek
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 64901
; LENGTH: 881
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-64901

Query Match      99.4%; Score 1748; DB 22; Length 881;
Best Local Similarity 99.1%; Pred. No. 2.6e-171;
Matches 335; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY      1 MAALAOEDGWTGGOVLVYVNSAGDAIGLOPDARGVATSLGNERLFVYVNPQEVHELIPHP 60
```



Db	544	MAALAOEDGWTGQVLVKVNSAGDAIGLOPDAGVATSLGLNERLPVNPQEVHELIPHP	603
QY	61	DOLGPTVGSABEGLDVSAKDLAQLTBDHMSLFNSIHQVELIHYVLGPOHLRDVTTANLE	120
Db	604	DOLGPTVGSABEGLDVSAKDLAQLTBDHMSLFNSIHQVELIHYVLGPOHLRDVTTANLE	663
QY	121	RFMRFPNELQYVWATELCLCPVGPRAQLLRKFIKLAHLKEQKNLNSFFAVMFGLSNSA	180
Db	664	RFMRFPNELQYVWATELCLCPVGPRAQLLRKFIKLAHLKEQKNVNSFFAVMFGLSNSP	723
QY	181	ISRLAHTWERLPHKVKKLYSALERLDPGSMNHRVYRLALAKLSPVYIPFMPLLLKDMTEI	240
Db	724	ISRLAHTWERLPHKVKKLYSALERLDPGSMNHRVYRLALAKLSPVYIPFMPLLLKDMTEI	783
QY	241	HEGNHTLVENLINFEEKRMARARMLHCRSHNPVPLSPLRSRVSHLHEDSOVARISTC	300
Db	784	HEGNHTLVENLINFEEKRMARARMLHCRSHNPVPLSPLRSRVSHLHEDSOVARISTC	843
QY	301	SEOSLSTRSPASTWAVVQOLKVIDNQRELSRLSRELEP	338
Db	844	SEOSLSTRSPASTWAVVQOLKVIDNQRELSRLSRELEP	881

Search completed: October 7, 2003, 07:23:37  
Job time : 404 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 07:15:33 : Search time 66 Seconds  
(without alignments)  
810.242 Million cell updates/sec

Title: US-09-856-679-2  
Perfect score: 1759  
Sequence: 1 MAALAEDEGWTGQVLVKVN.....QLKVIDNQRELRLSHELPE 338

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues  
Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1408	80.0	270	10	US-09-911-826A-8
2	419.5	23.8	246	10	US-09-940-836A-6
3	418.5	23.8	238	10	US-09-922-199A-9
4	409	23.3	1499	10	US-09-911-826A-2
5	405.5	23.1	261	9	US-09-822-635-4
6	405.5	23.1	261	16	US-10-176-306-32
7	396.5	22.5	307	10	US-09-911-826A-7
8	246.5	14.0	227	10	US-09-922-199A-5
9	246.5	14.0	227	16	US-10-176-306-31
10	238	13.5	472	10	US-09-940-836A-2
11	237	13.5	473	5	US-10-275-173-2
12	229	13.0	59	9	US-09-864-761-39774
13	221	12.6	591	12	US-09-805-455-7
14	207.5	11.8	782	15	US-10-177-293-399
15	204	11.6	244	10	US-09-911-826A-9

16	196	11.1	243	10	US-09-911-826A-11	Sequence 11, App1
17	194	11.0	709	14	US-10-118-328-4	Sequence 4, App1
18	193	11.0	684	10	US-09-765-298A-18	Sequence 18, App1
19	191.5	10.9	201	15	US-10-103-313-420	Sequence 420, App
20	190.5	10.8	581	12	US-09-805-455-8	Sequence 8, App1
21	190	10.8	428	15	US-10-157-031-139	Sequence 139, App
22	186	10.6	164	15	US-10-103-313-543	Sequence 543, App
23	182	10.3	699	15	US-10-233-131-34	Sequence 34, App1
24	182	10.3	710	14	US-10-118-328-2	Sequence 2, App1
25	182	10.3	710	15	US-10-060-990-3	Sequence 3, App1
26	182	10.3	715	15	US-10-233-131-33	Sequence 33, App1
27	178	10.1	1589	10	US-09-801-368-64	Sequence 64, App1
28	177.5	10.1	249	10	US-09-911-826A-10	Sequence 10, App1
29	175.5	10.0	777	15	US-10-205-823-345	Sequence 345, App1
30	170	9.7	169	10	US-09-940-836A-4	Sequence 4, App1
31	169.5	9.6	608	12	US-09-805-455-6	Sequence 6, App1
32	166.5	9.5	609	12	US-09-805-455-4	Sequence 4, App1
33	166.5	9.5	615	12	US-09-805-455-2	Sequence 2, App1
34	166.5	9.5	664	12	US-09-805-455-5	Sequence 5, App1
35	166	9.4	528	9	US-09-925-297-806	Sequence 806, App
36	162.5	9.2	926	14	US-10-108-605-143	Sequence 143, App
37	161.5	9.2	673	16	US-10-176-306-2	Sequence 2, App1
38	157.5	9.0	142	10	US-09-764-868-1023	Sequence 1023, App
39	157.5	9.0	142	15	US-10-103-313-429	Sequence 429, App
40	155	8.8	158	15	US-10-103-313-404	Sequence 404, App
41	150	8.5	677	15	US-10-103-313-384	Sequence 384, App
42	150	8.5	902	10	US-09-922-199A-2	Sequence 2, App1
43	148.5	8.4	725	14	US-10-118-328-5	Sequence 5, App1
44	143	8.1	141	15	US-10-275-173-4	Sequence 4, App1
45	132.5	7.5	60	9	US-09-864-761-44037	Sequence 44037, A

#### ALIGNMENTS

RESULT 1  
US-09-911-826A-8  
Sequence 8, Application US/09911826A  
Patient No. US20020143164A1  
GENERAL INFORMATION:  
APPLICANT: Rotin, Daniela and Pham, Nam  
TITLE OF INVENTION: RAS Activator Nucleic Acid Molecules, Polypeptides and  
FILE REFERENCE: DDW-5001-US  
CURRENT APPLICATION NUMBER: US/09/911, 826A  
CURRENT FILING DATE: 2002-02-26  
PRIOR APPLICATION NUMBER: PCT/CA00/00042  
PRIOR FILING DATE: 2000-01-20  
PRIOR APPLICATION NUMBER: 2,259, 830  
PRIOR FILING DATE: 1999-01-20  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 8  
LENGTH: 270  
TYPE: PRT  
ORGANISM: Homo sapiens

Query Match	Score	80.0%	1408	DB 10;	Length 270;
Best Local Similarity	100.0%	Pred. No. 1.1e-138;			
Matches 270;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	69	SAGGLDLSAKDLAAGLTDDHDSLSPNSIHQVELIHYVLGPOHLRDVTANTLERPMRFNE	128		
DB	1	SAGGLDLSAKDLAAGLTDDHDSLSPNSIHQVELIHYVLGPOHLRDVTANTLERPMRFNE	60		
QY	129	LQYVATLCLCPVGPRAQLLRKFLKLAHLKEOKNLSPFAVMGLSATSRLAHTW	188		
DB	61	LQYVATLCLCPVGPRAQLLRKFLKLAHLKEOKNLSPFAVMGLSATSRLAHTW	120		
QY	189	ERLPHKVRKLYSALERLDPSSMNRKRYRLALAKLSPVYFPMLLKDMFTIEGNTLV	248		
DB	121	ERLPHKVRKLYSALERLDPSSMNRKRYRLALAKLSPVYFPMLLKDMFTIEGNTLV	180		

```

QY      249  ENLINEKKRMMARAAARMILHCRSHNPVPLSPILSGRVSHLHEDSOVARISCSPOSILSTR 308
Db      181  ENLINEKKRMMARAAARMILHCRSHNPVPLSPILSGRVSHLHEDSOVARISCSPOSILSTR 240
QY      309  SPASTWAVYOOLKVIDNORELSRSLRELEP 338
Db      241  SPASTWAVYOOLKVIDNORELSRSLRELEP 270

RESULT 2
US-09-940-836A-6
: Sequence 6, Application US/09940836A
: Patent No. US20020146800A1
: GENERAL INFORMATION:
: APPLICANT: Millennium Pharmaceuticals, Inc.
: APPLICANT: Curtis, Rory
: TITLE OF INVENTION: 48921, A NOVEL HUMAN GTP RELEASING
: FILE REFERENCE: FACTOR AND USES THEREFOR
: CURRENT APPLICATION NUMBER: US/09/940,836A
: CURRENT FILING DATE: 2001-08-27
: PRIOR APPLICATION NUMBER: US 60/228,760
: PRIOR FILING DATE: 2000-08-30
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 6
: LENGTH: 246
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Consensus amino acid sequence
US-09-940-836A-6

Query Match      23.8%; Score 419.5; DB 10; Length 246;
Best Local Similarity 43.3%; Pred.No. 2.2e-35;
Matches 91; Conservative 40; Mismatches 58; Indels 21; Gaps 5

QY      73  LD- -VSARDLAGQLTDHDMSLFNSIHQVELIHV-----LGPQHLRDVYTANLEPRM 124
Db      8  LDLDLHDPBELAEQLLDLDFELFKIEPSDELGYWMSNREKKGENT----SPRLKEFTIG 63
QY      125  RNELOYWATLCLCPVPGPAQLLRKFIKLAHLKQKLNSEFPAVMFGLSNATSIRL 184
Db      64  RENNISYWATLLETLSSEKFOQAKIIEKFIKVAOHCRELNFNLSMAIVSGLNSSSIYRL 123
QY      185  AHWHELPKVKRYKTSALERLIDPSMNRVRYRLAKLS-----PPVLPFMDLLIKOM 237
Db      124  KATWEVVPPEYTKLKEELSELMDPSNNKNTRELKSCGTGVHSHQPCIPFVGVIYLDK 183
QY      238  TFIHESGNTLVE- -NLINFEKKRMARARAR 265
Db      184  TFIHESGNDPYLDNTNLINFEKKRMIAKILIR 213

RESULT 3
US-09-922-199A-9
: Sequence 9, Application US/09922199A
: Publication No. US20020187138A1
: GENERAL INFORMATION:
: APPLICANT: Meyers, Rachel
: TITLE OF INVENTION: 15368, A NOVEL HUMAN GTP-RELEASING
: FILE REFERENCE: FACTOR FAMILY MEMBER AND USES THEREFOR
: CURRENT APPLICATION NUMBER: US/09/922,199A
: CURRENT FILING DATE: 2001-08-02
: PRIOR APPLICATION NUMBER: 60/222,622
: PRIOR FILING DATE: 2000-08-02
: NUMBER OF SEQ ID NOS: 28
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 238
: TYPE: PRT

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-922-199A-9

Query Match      23.8%; Score 418.5; DB 10; Length 238;
Best Local Similarity 44.1%; Pred. No. 2,7e-35;
Matches 89; Conservative 38; Mismatches 56; Indels 19; Gaps 4

QY      79  KDLAQLDTHDMSLFNSIHQVELHYV-----IGPOLRVDTNTERPMRFNELYW 132
Db      7  EELAEQLTLDFELFKTEPSECCTGYVSNREKKCKEMD-----SNLEKFIORFNISYW 62

QY      133 VATELCPCVPGPRAQLLRKFIKLAHLKEQKINSFPAVFGLSNSAISRLAHTWERLP 192
Db      63  VATELISEKPOQRAKITEKFTKVAHQREINNFSLMAIVSGLSNSISYRLKTKWEKVP 122

QY      193  HKVAKYLSALERLDPSNNHNYVRLAKLS-----PVIYFPPLLKDMFTFHEGHN 245
Db      123  KETKLEFEELSDLPSSNNYNYRELKSCGTFVNHSDPCPIPLGVYKLDLFTFHEGNP 182

QY      246  TLVE--NLINFEKRRMARAR 265
Db      183  DYLDNTNLINFEKRRMIAKIR 204

RESULT 4
US-09-911-826A-2
; Sequence 2, Application US/09911826A
; Patent No. US20020143164A1
; GENERAL INFORMATION:
; APPLICANT: Rotin, Daniela and Pham, Nam
; TITLE OF INVENTION: RAS Activator Nucleic Acid Molecules, Polypeptides and
; FILE OF INVENTION: Methods of Use
; FILE REFERENCE: DDW-5001-US
; CURRENT APPLICATION NUMBER: US/09/911.826A
; CURRENT FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: PCT/CA00/00042
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: 2,259,830
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1499
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-911-826A-2

Query Match      23.3%; Score 409; DB 10; Length 1499;
Best Local Similarity 30.9%; Pred. No. 4,2e-33;
Matches 107; Conservative 57; Mismatches 122; Indels 60; Gaps 5

QY      35  VATSLGINERLFEVVPQEVHELIPHPDLAGPTVGSAGELDIVSAKDLAGQLTDHWSLFN 94
Db      675  LADRIQISGRYYLTKNNMETETLCDEDDAQELLRSQSLDLSLTSEVATQLSMRFFELR 734

QY      95  SIHQVELIHYVGGQHLRDYTT-ANLEFPMRFNELYWVATELCPCVPGPRAQLLRKF 153
Db      735  NIEPTXYIDFLE--KLRSKTSKANLKFEFEVINOETFWVASSELLRETNOIKRMKIKHF 791

QY      154  IKLAHLKEQKINSFPAVFGLSNSAISRLAHTWERLPHVKKYLSALERLDPSSNHR 213
Db      792  IKIALHRECKENFSMFAIISGLMLAPARLRTIWEKPKPKNYEKLFOQLDQLDFDSRMA 851

QY      214  VYRLAL--AKISPVIDFPMPLLKDMFTFHEGHNFTLVENLINFEKRRMARARMLHCR 271
Db      852  KRYVNLNSQNLQPIPIELFPIYKIDLFTFHEGNSKSDGLVNFELKRIAKEIRIVGMA 911

QY      272  SHNVPISPLRSR-----VSHLHDS 292
Db      912  SVNDDPALMFTKTRKKKRRSLGSLGSGSTNATVLDVAQTGGHKRYRNSSTLNAKKLYEDA 971

```

OY 293 QVARISTCSQSLSSTRSPASTWAVYQOLKVIDNORELSRLSRELEP 338  
Db 972 QMAR-----KVQYLSNLELEMBEESQTLSTLQCEP 1002

## RESULT 5

US-09-822-635-4  
; Sequence 4, Application US/09822635  
; Patent No. US20010039331A1  
; GENERAL INFORMATION:  
; APPLICANT: Hunter, John J.  
; APPLICANT: Meyers, Rachel A.  
; TITLE OF INVENTION: 16836, A NOVEL HUMAN PHOSPHOLIPASE C AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: 10448-035001  
; CURRENT APPLICATION NUMBER: US/09/822,635  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: US 60/193,921  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 261  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: consensus sequence  
US-09-822-635-4

## Query Match

23.1%: Score 405.5; DB 9; Length 261;  
Best Local Similarity 37.6%; Pred. No. 7e-34;

Matches 106; Conservative 41; Mismatches 92; Indels 43; Gaps 8;

OY 75 LVSAKDLAAGLTLDHMSLFSIHQVELIHVLPQHLD---VTTALERFMRFNLOY 131  
Db 3 LDPDELEAQLTLDLDFELFKRIEPELIGSVGKRSKSPSPPLAPQULEAFIERFNEYSN 62  
OY 132 WVAETELCLCPV--PGPRAQLRKFKIKLAHLKEQKLNSEFAVNGLSNALSRLATWE 189  
Db 63 WVAETELKQTTLKPKRAEVLKFIKAKKRELNNSLMATVSALSSPSISRLKTWE 122  
OY 190 RLPHKVKLYSALERLDP--WNHRYRLAL-----AKLSPYIPMPPLKMDPTIHE 242  
Db 123 KLSKYKKLEPELELDPSEERFNKRYREALSCNKSPPNQPVPPLGYLKDLPIDE 182  
OY 243 GNHITVEN---LINFEMKMMARAAMLHCRSH-NPVPLSPRSRVSHLEDQVARI 297  
Db 183 GNPDELENGTKGLVNFERRKRIKILREIROLQACOPYNLKPRNDIQELLRAS----- 237  
OY 298 STCSQSLSSTRSPASTWAVYQOLKVI-DNORELSRLSRELEP 338  
Db 238 -----RPLEVLPDEEDELVELSLRIEP 259

## RESULT 6

US-10-176-306-32  
; Sequence 32, Application US/10176306  
; Publication No. US20030130465A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel A.  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Gluckmann, Maria Alexandra  
; APPLICANT: Bandaru, Rajasekhar  
; APPLICANT: Kapeller-Libermann, Rosana  
; TITLE OF INVENTION: NOVEL HUMAN GENES AND METHODS OF USE THEREOF  
; FILE REFERENCE: 10448-195001  
; CURRENT APPLICATION NUMBER: US/10/176,306  
; PRIOR APPLICATION NUMBER: 2002-06-20  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: PCT/US01/45291  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/248,362

; PRIOR FILING DATE: 2000-11-14  
; PRIOR APPLICATION NUMBER: 60/248,331  
; PRIOR FILING DATE: 2000-11-14  
; PRIOR APPLICATION NUMBER: 60/248,365  
; PRIOR FILING DATE: 2000-11-14  
; PRIOR APPLICATION NUMBER: 60/250,077  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 60/250,327  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 60/250,176  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 10/023,617  
; PRIOR FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: PCT/US01/49416  
; PRIOR FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: 60/256,249  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: 60/256,405  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: 10/083,248  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: PCT/US01/46717  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: 60/242,324  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/242,518  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/241,989  
; PRIOR FILING DATE: 2000-10-20  
; NUMBER OF SEQ ID NOS: 86  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 32  
; LENGTH: 261  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: consensus sequence  
US-10-176-306-32

Query Match 23.1%: Score 405.5; DB 16; Length 261;  
Best Local Similarity 37.6%; Pred. No. 7e-34;

Matches 106; Conservative 41; Mismatches 92; Indels 43; Gaps 8;

OY 75 LVSAKDLAAGLTLDHMSLFSIHQVELIHVLPQHLD---VTTALERFMRFNLOY 131  
Db 3 LDPDELEAQLTLDLDFELFKRIEPELIGSVGKRSKSPSPPLAPQULEAFIERFNEYSN 62  
OY 132 WVAETELCLCPV--PGPRAQLRKFKIKLAHLKEQKLNSEFAVNGLSNALSRLATWE 189  
Db 63 WVAETELKQTTLKPKRAEVLKFIKAKKRELNNSLMATVSALSSPSISRLKTWE 122  
OY 190 RLPHKVKLYSALERLDP--WNHRYRLAL-----AKLSPYIPMPPLKMDPTIHE 242  
Db 123 KLSKYKKLEPELELDPSEERFNKRYREALSCNKSPPNQPVPPLGYLKDLPIDE 182  
OY 243 GNHITVEN---LINFEMKMMARAAMLHCRSH-NPVPLSPRSRVSHLEDQVARI 297  
Db 183 GNPDELENGTKGLVNFERRKRIKILREIROLQACOPYNLKPRNDIQELLRAS----- 237  
OY 298 STCSQSLSSTRSPASTWAVYQOLKVI-DNORELSRLSRELEP 338  
Db 238 -----RPLEVLPDEEDELVELSLRIEP 259

## RESULT 7

US-09-911-826A-7  
; Sequence 7, Application US/09911826A  
; Patent No. US20020143164A1  
; GENERAL INFORMATION:  
; APPLICANT: Rotin, Daniela and Pham, Nam  
; TITLE OF INVENTION: RAS Activator Nucleic Acid Molecules, Polypeptides and  
; FILE REFERENCE: DDM-5001-US

```
; CURRENT APPLICATION NUMBER: US/09/911,826A
; CURRENT FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: PCT/CA00/00042
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: 2,259,830
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-911-826A-7

Query Match
Best Local Similarity 22.5%; Score 396.5; DB 10; Length 307;
Best Local Similarity 34.3%; Pred. No. 7,8e-33;
Matches 103; Conservative 54; Mismatches 102; Indels 41; Gaps 9;

QY 76 VSAKDLAAGLTLDHMSLFNSIHQVELIHVLPQHLD-VTTANLERFKRRENELOYWA 134
DB 8 LNAVELAIQLTLD--FANFROISTEVEYDELFEELRSRYGVPMLSKFAELVNRKMFVW 64
QY 135 TELCLCPVPGPRAQLLRKFIKLAHLKEQKNLSFPAVMEGLSNAISRLAHTWERTPRK 194
DB 65 SEICAEHNIVRRMKIVKQIKIARHCKEGRNFSMAISGLHGAVSLRQTEKTPSK 124
QY 195 VKKLSALERLDPSNMHRVYR-LALAKL-SPPVIEPMLLKDMFTIEGNNHTLVENL 251
DB 125 YORLTNDLMDPSNMKSKYQLVSAELIAOHPIPEYPIVKDLFTHLGNDTRVDGL 184
QY 252 INFEKRMARARAHMHCRSHNPVPL-----SPLR-----SRVSHLEDQSVARI 297
DB 185 VNEKRLMAKRVRLTHMCS-SPYDLISLELKQSPSNALFSLNOMASOSNAAGTV 243
QY 298 -----STCSQSLSSTRSP-----ASTWAVYQOLKVIDNQRELSRLSRELEP 338
DB 244 IAAAGATIKRRKKSTAPNPKKMEEQMVRKAYLNSLKLISDEDLNKFSLCEP 303

RESULT 8
US-09-922-199A-5
; Sequence 5, Application US/09922199A
; Publication No. US20020187138A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 15368, A NOVEL HUMAN GTP-RELEASING
; TITLE OF INVENTION: FACTOR FAMILY MEMBER AND USES THEREFOR
; FILE REFERENCE: 381552002500
; CURRENT APPLICATION NUMBER: US/09/922,199A
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/222,622
; PRIOR FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-922-199A-5

Query Match
Best Local Similarity 14.0%; Score 246.5; DB 10; Length 227;
Best Local Similarity 30.2%; Pred. No. 2,4e-17;
Matches 68; Conservative 39; Mismatches 77; Indels 41; Gaps 6;

QY 75 LVSADLAGLTLDHMSLFNSIHQVELIHVYL---GPO-----HLDVTTANLERPMR 124
DB 2 LLDPLELAKQLTLEHLEFKITTPRECIGSKMKHGPQWCKSKLKGEEERSPNIDKTIK 61
QY 125 RNELOQYWA-----TELCLCPVPGPRAQLLRKFIKLAHLKEQKNLSFPAVMEGL 173
DB 62 NFQNLTFVSEIILKQEPPTNPEFSKKYDPKRAEYIOKFIQVADHCHLNFNNSLAIIT 121
```

```
QY 174 FGLNSAISRLAHTWERTPRKVRKL-YSALERLDPSNMHRVYRLALAKLSP----- 224
DB 122 SALYSSPIYRLKTKTEKVPQSLKLFEEELNKLMSDRNFSYRELKLSIGLKPNNLDD 181
QY 225 -----PVIEPMLLKDMFTIEGNNHTLVE-----NLINFEKMR 258
DB 182 DFVRAPCVPEFGVYLSDLTFLEEGNDYLENKRINTNLVNFESKR 226

RESULT 9
US-10-176-306-31
; Sequence 31, Application US/10176306
; Publication No. US20030130485A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: NOVEL HUMAN GENES AND METHODS OF USE THEREOF
; FILE REFERENCE: 10448-195001
; CURRENT APPLICATION NUMBER: US/10/176,306
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 10/001,137
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: PCT/US01/45291
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/248,362
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/248,331
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/248,365
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/250,077
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,327
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,176
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 10/023,617
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: PCT/US01/49416
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,249
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/256,405
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 10/083,248
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/US01/46717
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/242,324
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,518
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/241,989
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-10-176-306-31

Query Match
Best Local Similarity 14.0%; Score 246.5; DB 16; Length 227;
Best Local Similarity 30.2%; Pred. No. 2,4e-17;
Matches 68; Conservative 39; Mismatches 77; Indels 41; Gaps 6;

QY 75 LVSADLAGLTLDHMSLFNSIHQVELIHVYL---GPO-----HLDVTTANLERPMR 124
```



```

; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39774
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004241.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.2
; OTHER INFORMATION: EST_HUMAN HIT: BE379558.1, EVALUATE 5.00e-19
; OTHER INFORMATION: SWISSPROT HIT: P48308, EVALUATE 6.00e+00
US-09-864-761-39774
```

```

Query Match          13.0%; Score 229; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.1e-16;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 57 IHPDQIGPTVGSABGLDVLVSADLAGQLTDHDMSLFNSTHOV 99
DB 1 IHPDQIGPTVGSABGLDVLVSADLAGQLTDHDMSLFNSTHOV 43
```

```

RESULT 13
US-09-805-455-7
; Sequence 7, Application US/09805455
; Publication No. US20030166203A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weimin
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE REFERENCE: C1001165
; CURRENT APPLICATION NUMBER: US/09/805,455
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Human
US-09-805-455-7
```

```

Query Match          12.6%; Score 221; DB 12; Length 591;
Best Local Similarity 24.8%; Pred. No. 4.7e-14;
Matches 84; Conservative 55; Mismatches 122; Indels 78; Gaps 10;
```

```

QY 26 ISLQDPARCVATSLGNERLFFVNPQEVHELIPHPDQIGPTVGSABGLDVLVSADLAGQL 85
DB 95 IHWTEEFREVASOLGYEKHVSIDISS---IPSYDMRMRYVQRKK---VSKKGACLL 146
QY 86 THHDSLFNSIHQVELIHVLPQH--LDDVT-----TANLERFMRRFE 128
DB 147 FHH-----LEPIELAEHLFLEHFSFRISFTDYQSYVHIGCLENNPTLERSTALFNG 199
QY 129 LQYVATELCLCPVGPQRQLRKFKTLAHLKEQKNLSFPAVMGSLNSAISRLAHW 188
```

```

DB 200 ISKWOVLAMVSKPTPOQRAEVITKFTINAKKLLQLKNEFTLMAVGGLSHSSISRLKETH 259
QY 189 ERLPHKRVKLYSALERLDPSPWNRVRYRLAKLSPVPIPFMPLLLKDMFTFHEGHTLV 248
DB 260 SHLSSEVTKNMNMETLVSNGVYKRAFPDCCDFKPIILGVHLKDLIAVH-----VI 314
QY 249 -----ENLINFKKMMMAARAAMLHCHSHNPVLS---PLRSKRVSHLHDSQVARIKT 299
DB 315 FPDWTEENKYNVIKMHQLS-----VTLSELVSLQNASHLPEPMDLITLLT 360
QY 300 CSEQSLSTRSPASTWAVYQQLKVIDNORELSRLEP 338
DB 361 LS-----LDLYHTEDDIYKLIVLEP 381
```

```

RESULT 14
US-10-177-293-399
; Sequence 399, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatte, Karen
; APPLICANT: Zhao, Xumel
; APPLICANT: Ganavarpu, Manjula
; APPLICANT: Kamalakar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Puzsrai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 399
; LENGTH: 782
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-399
```

```

Query Match          11.8%; Score 207.5; DB 15; Length 782;
Best Local Similarity 22.7%; Pred. No. 1.9e-12;
Matches 71; Conservative 50; Mismatches 101; Indels 91; Gaps 6;
```

```

QY 51 QEVHELIPHPDQIGPTVGSABGLD-----VSADLAGQLTDHDMSLFNSTHOVEL 101
DB 136 EEFQELV-----KAKGEELHCRLLDITTOJNARDWSKRLQRIKSNFSKRRKVS 184
QY 102 IHVYGLPOHLRDVTT-----ANLERFMRRNELEQYVVA 134
```



Db 185 LFDHLEPEELSEHLTYLEFESFRRISFSDYQNYLVNSCVKENPTMERSIALCNGISQWQ 244  
QY 135 TELCLCPVGPRAQLLRKFIKLAHLKEOKNLNSFPAVMGCLNSAISRLAHTERLPHK 194  
Db 245 LMYLSRPTPOLRAEVFTKFIQVAKLQLOLNFNTLMAVIGLCHSSISRLKETSSHPHE 304  
QY 195 VRKLVSALERLDPSSMNRVRYRLAKLSPPVIPPMLLKDMTFIHEG----- 243  
Db 305 INKVLGEMTELLSSRRYDNRAYGECTDEKIPILGVHLKDLISLYEAMPDYLGDKVN 364  
QY 244 -----NHTLVENLINFEMRMARARMLHRC-----RS 272  
Db 365 VHKLLALYNH--ISELVQLQEVAPPLEANKDVLHLTLSDLYYTEDEIYELSYAREPRN 422  
QY 273 HNPVPLSPLRGRV 285  
Db 423 HRAPPLTPSKRPV 435

## RESULT 15

US-09-911-826A-9  
: Sequence 9, Application US/09911826A  
: Patent No. US20020143164A1  
: GENERAL INFORMATION:  
: APPLICANT: Rotin, Daniela and Pham, Nam  
: TITLE OF INVENTION: RAS Activator Nucleic Acid Molecules, Polypeptides and  
: TITLE OF INVENTION: Methods of Use  
: FILE REFERENCE: DDM-5001-US  
: CURRENT APPLICATION NUMBER: US/09/911,826A  
: CURRENT FILING DATE: 2002-02-26  
: PRIOR APPLICATION NUMBER: PCT/CA00/00042  
: PRIOR FILING DATE: 2000-01-20  
: PRIOR APPLICATION NUMBER: 2,259,830  
: PRIOR FILING DATE: 1999-01-20  
: NUMBER OF SEQ ID NOS: 27  
: SOFTWARE: PatentIn Ver. 2.1  
: SEQ ID NO 9  
: LENGTH: 244  
: TYPE: PRT  
: ORGANISM: Mus musculus  
US-09-911-826A-9

Query Match 11.6%; Score 204; DB 10; Length 244;  
Best Local Similarity 25.8%; Pred. No. 7.4e-13;  
Matches 60; Conservative 44; Mismatches 115; Indels 14; Gaps 4;

QY 70 AEGIDLVSADLAQQLDHDMSLENSIHVELIHYVLGPOHL--RDVTTANLERFMRRF 126  
Db 2 AECFETLSAMELAQITLDDHIVFRSIPYEEF---LGQGMKLDKKNERTPYIMKTSQHF 57  
QY 127 NELQYVATELCPCVPGRPAQLLRKFIKLAHLKEOKNLNSFPAVMGCLNSAISRLAH 186  
Db 58 NEMSNLVAASQIMNADISSRPALEKRVAVADICRCHNNGVLEITSALNRSPIYRLKK 117  
QY 187 TWERLPHKVRKLYSALERLDPSSMNRVRYRLAKLSPPVIPPMLLKDMTFIHEGNNHT 246  
Db 118 TMAKVSQOTKALMDKLOKTVSSEGRPKMLRETLKNCNRPAPVYLGWYLTDLAFLIEGTPN 177  
QY 247 LV-ENLINFEMRMARARMLHRC-----SHNPVPLSPLRGRVSHLHEDS 292  
Db 178 FTEGLVNFSKRMISIRIROIROFOOTAVRIDOOPKVIQYLDLKDALVIDEDS 230

Search completed: October 7, 2003, 07:25:21  
Job time : 67 secs

**THIS PAGE BLANK (USPTO)**







```

C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999
C:Accession: G29083
R:Shou, C.; Farnsworth, C.L.; Neel, B.G.; Feig, L.A.
A>Title: Molecular cloning of cDNAs encoding a guanine-nucleotide-releasing factor for F
F:240-426/Domain: CDC24 homology <CD24>
A:Accession: G29083
A:Molecule type: mRNA
A:Residues: 1-1244 <SHO>
A:CROSS-references: EMBL:X67241; NID:g57664; PIDN:CAA7666.1; PID:g57665
A>Note: The authors translated the codon GAG for residues 135 and 137 as Glu
C:Superfamily: CDC25-type guanine nucleotide exchange activator homology; CDC24 homology
F:1005-1241/Domain: CDC25-type guanine nucleotide exchange activator homology <SOS>

Query Match          11.7%; Score 206; DB 2; Length 1244;
Best Local Similarity 23.9%; Pred. No. 1e-08;
Matches      68; Conservative 56; Mismatches 144; Indels 16; Gaps 6;

OY    28  LQPDARGVATSL--GLNERLFVNVPNOEHNELPHRDPLGPTVSGAEGDLVASAKDLGOL 85
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db     963  LLPEBKRAAIIMMTLVQETETENHSMDELLMTEGV-----KTEFFENHSAMETIRQL 1017

OY    86  TDHWSLFNSIHVELIHYVLGPQH---RDVTANLERFERRNEDLYWATELCIPV 142
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db     1018 TLDELHVFKSPIREEF----FGQGMMKADKKERFPRTYMKTRHHNHISNLASIELRNEE 1073

OY    143  PGPRAAOLLRRIKTKLAHLKEQNLSFFAVFGISNAISLATWERLPKHVKKLVSAL 202
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db     1074 VSAARASTIEKWVAADVDCRCCLAHNYNAVLEITSSINRSAIRFLKKTMLKVSKOYSLFDKL 1133

OY    203  ERLLDPSMNHRVRYALAKSPPIYPMPRLCLKMTEFIHEGNHTLVEN-LINEKRMMA 261
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db     1134 QKIVSDSGREFNLRETFRNCDPFCVPYLGMVLTDLAELEGSTPYTEDGLVNFSGKMIMS 1193

OY    262  RAARMILHCGRSHNPVPLSPILSRVSHLHEDSQVARISTCSQS 305
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db     1194 HIREIRIFO-QTTYKIEPQPKYTGYLVDEFVLDDSELYEASL 1236

RESULT 9
T42726
guanine nucleotide release/exchange factor Ras-GRF2 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Aug-2000
C:Accession: T42726
R:Fam, N.P.; Fan, W.; Zhang, L.; Chen, H.; Moran, M.F.
A>Title: Cloning and characterization of Ras-GRF2, a novel guanine nucleotide exchange f
A:Reference number: Z22245; MUID:97184464; PMID:9032266
A:Accession: T42726
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1189 <FAM>
A:CROSS-references: EMBL:U67326; NID:g1655940; PID:g1655941; PIDN:AAC53058.1
A:Experimental source: brain
C:Genetics:
A:Map position: 13
C:Superfamily: CDC25-type guanine nucleotide exchange activator homology
F:950-1186/Domain: CDC25-type guanine nucleotide exchange activator homology <SOS>

Query Match          11.6%; Score 204; DB 2; Length 1189;
Best Local Similarity 25.8%; Pred. No. 1.4e-08;
Matches      60; Conservative 44; Mismatches 115; Indels 14; Gaps 4;

OY    70  AEGDIIVASDKLDAGQLTDHWSLFNSIHVELIHYVLGPQH---RDVTANLERFERRP 126
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db     947  AACEFTLSAMELARQITLDHIYFRSPIREBF----LGCGMMKDKKERFPRTYMKTSQH 1002

OY    127  NELQYWATELCICPVPGPRAQLRKIKLAHLKEKNLNSFFAVFGISNASISLAH 186
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db     1003 NEKSNILAASIMNVADISSPPNALTEKWVAADVDCRCCLAHNGVLEITSALNRSPITYLRK 1062

```

```

Oy      187 TWELPHKVRKLYALERLDPDSNMHRYRLAKLSPVPEFMPILDKDWTFTHEGNHT 246
      188 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1063 TMAVSNQTKRLMDKLOKTVSSEGRFKNLRRLTNCNCPNPAVPIGMYITDLAFLTEEGPN 1122

Oy      247 LV-ENLINFEMKMMARAARMLHCCR-----SHNPVPLSPRSRVSHLED 292
      1123 FTKEGLVNFSSKRMISHIRFIRFOQTAVRIDQPKVIQYLIDKALVIDEDS 1175

RESULT 10
S28407
guanine nucleotide-exchange activator CDC25 homolog - mouse
C:Species: Mus musculus (house mouse)
C:Date: 17-Apr-1993 #sequence.revision 17-Apr-1993 #text_change 05-Nov-1999
C:Accession: S28407; S22693; B46199; S20730
R:Cen, H.; Papageorge, A.G.; Zippel, R.; Lowy, D.R.; Zhang, K.
EMBO J. 11, 4007-4015, 1992
A:Title: Isolation of multiple mouse CDNA's with coding homology to Saccharomyces cerevisiae guanine nucleotide-exchange factor CDC25
A:Reference number: S28407; MUID:93010996; PMID:1396590
A:Accession: S28407
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1260 <EN>
R:Mattegan, E.; Vanoni, M.; Zippel, R.; Coccetti, P.; Brambilla, R.; Ferrari, C.; St
EMBO J. 11, 2151-2157, 1992
A:Title: Cloning by functional complementation of a mouse cDNA encoding a homologue c
A:Reference number: S22693; MUID:92289680; PMID:1376246
A:Accession: S22693
A:Molecule type: mRNA
A:Residues: 789-1260 <MAR>
A:Cross-references: EMBL:X59868; NID:950357; PIDN:CAA42525.1; PID:950358
R:Met, W.; Mosteller, R.D.; Sanyal, P.; Gonzales, E.; McKinney, D.; Dasgupta, C.; Li,
Proc. Natl. Acad. Sci. U.S.A. 89, 7100-7104, 1992
A:Title: Identification of a mammalian gene structurally and functionally related to
A:Reference number: A46199; MUID:92357779; PMID:1379731
A:Accession: B46199
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1029-1030, 'D', 1032-1224 <MEI>
A:Experimental source: fetus
A:Note: Sequence extracted from NCBI backbone (NCBIN:111101, NCBIP:111102)
C:Superfamily: CDC25-type guanine nucleotide exchange activator homology; CDC24 homol
F:242-428/Domain: CDC24 homology <CD24>
F:1021-1257/Domain: CDC25-type guanine nucleotide exchange activator homology <SOS>

Query Match      11.5%; Score 201.5; DB: 2; Length 1260;
Best Local Similarity 25.1%; Pred. No. 2.4e-08;
Matches 58; Conservative 51; Mismatches 115; Indels 7; Gaps 4;

Oy      78 AKDLAGOLTDHMDLSPNSIHQVELI--HYVLGPOLHLDVYTANLERFMRPNELQYVAT 135
      1026 ALAETAEQTLTLDHLVFSISPEEFGGWMKAKEYER---TPRYMKTTHKFNHNSFIAS 1082
Db      1026 ALAETAEQTLTLDHLVFSISPEEFGGWMKAKEYER---TPRYMKTTHKFNHNSFIAS 1082

Oy      136 ELCLCPVGPRAQILRRFIKLAHLKQKLNINSEFAVFGLSNSAISRLAHTWERLPKHV 195
      1083 EIINEDISARASAIKFWVAADICRLHNNAVLEITSSINRSAIRLKTWKLKVSQQT 1142
Db      1083 EIINEDISARASAIKFWVAADICRLHNNAVLEITSSINRSAIRLKTWKLKVSQQT 1142

Oy      196 RKLYSALERLLDPSMNRRVRLALAKSPVPIPPMLLDMDPTIHGNGNTIVN-ILNF 254
      1143 KSLLDKLOKLVSSDGRFKNLRRESLNCDDPCPVIGMYITDLVDFIEGTPTNYEDGLVNF 1202
Db      1143 KSLLDKLOKLVSSDGRFKNLRRESLNCDDPCPVIGMYITDLVDFIEGTPTNYEDGLVNF 1202

Oy      255 EKMMEMARAARMLHCRSHNPVPLSPRSRSVSHLEDGQVARIStCEOSL 305
      1203 SKMMISHIRFIRFOQTAVRIDQPKVIQYLIDDESFMLEDSESL 1252
Db      1203 SKMMISHIRFIRFOQTAVRIDQPKVIQYLIDDESFMLEDSESL 1252

RESULT 11
A38985
nucleotide exchange factor CDC25 - human
N:Alternate names: Ras-specific guanine nucleotide-releasing factor
C:Species: Homo sapiens (man)
C:Date: 19-Jul-1996 #sequence.revision 19-Jul-1996 #text_change 05-Nov-1999

```

A:Accession: A38985; A46199; 158371  
R:Wel, W.; Brock, D.  
submitted to Genbank, December 1994  
A:Description: Cloning and analysis of the full length human cdc25 CDNA, a ras-specific  
A:Reference number: A38985  
A:Accession: A38985  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1275 <WE1>  
A:Cross-references: GB:I26584; NID:g4333719; PIDN:AA58417.1; PID:g4333720  
R:Wel, W.; Mosteller, R.D.; Sanyal, P.; Gonzales, E.; McKinney, D.; Dasgupta, C.; Li, P.  
Proc. Natl. Acad. Sci. U.S.A. 89, 7100-7104, 1992  
A>Title: Identification of a mammalian gene structurally and functionally related to the  
A:Reference number: A46199; MUID:9235779; PMID:1379731  
A:Accession: A46199  
A:Molecule type: mRNA  
A:Residues: 1047-1054, 'A', 1056-1112, 'G', 1114-1135, 'C', 1137-1275 <WE2>  
A:Experimental source: brain  
A:Note: sequence extracted from NCBI backbone (NCBIN:111098, NCBIPI:111099)  
R:Schweighoffer, F.; Faure, M.; Fath, I.; Chevallier-Multon, M.C.; Aplou, F.; Durrillaux  
Oncogene 8, 1477-1485, 1993  
A>Title: Identification of a human guanine nucleotide-releasing factor (H-GRF5) specific  
A:Reference number: 158371, MUID:93275641; PMID:7664828  
A:Accession: 158371  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 787-1275 <RES>  
A:Cross-references: GB:582035; NID:g386046; PIDN:AB26881.1; PID:g386047  
C:Superfamily: CDC25-type guanine nucleotide exchange activator homology; CDC24 homology  
E:240-46/Domain: CDC24 homology <CD24>  
E:1036-1277/Domain: CDC25-type guanine nucleotide exchange activator homology <SOS>

```

Query Match      11.3% Score 199; DB 2; Length 1275;
Best Local Similarity 27.8%; Pred. No. 4e-08;
Matches 55; Conservative 41; Mismatches 98; Indels 4; Gaps 3;

QY AEGDLVYSANDLACQLTDHMSLFNSIHQVELIHYVLGPOHL-RDVTANTLERPMRRFNE 128
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db AEPPNNSALEIAEQTLTLDHLVEFKKIPYEED--FGGGMMKLKEKNERTPIYMKTTHEND 1090
   |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY LQYVAATHELCCLPVGPGRAOLLRFETIKLAHLKEQQKNLNSEFAVMFGLSNSAIRLAWHT 168
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db ISNLIASEEIIIRNEDINARVASIEKMVAADVADICRLNHYNNAVLTTSSMNRSAIFELRKWT 1150
   :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY ERLPHKRYKLSALERLLDPSPMNHRVLRVALAKLSPVIPFPMLLKDMFTFIHGNGNTLV 248
   :::||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db LVSKQKKRALKLIDKQKLVSSEGGRFNLEREALKNDPCPVPLGMYLTIADIAIEBGTENYT 1210
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 249 EN-LINEEKKRMMAARAAR 265
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db EDGLVFNFSGMKRMISHITR 1228
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 12
SDC356
CDC25 protein homolog - yeast (Candida albicans)
S:Species: Candida albicans
C:Date: 20-Feb-1995 #sequence_rev1sion 20-Feb-1995 #text_change 29-Oct-1999
C:Accession: S30356
R:Goldberg, D.; Merbach, I.; Gross, E.; Levitzki, A.; Shmehen, G.
Eur. J. Biochem. 213, 195-204, 1993
A>Title: A Candida albicans homolog of CDC25 is functional in Saccharomyces cerevisiae.
A:Reference number: S30356; WUID:93238685; PMID:8477693
A:Accession: S30356
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1333 <COL>
A:Cross-references: EMBL:M94160; NID:g170834; PID:AAA3439.1; PID:g170835
C:Superfamily: SH3 homology; CDC25-type guanine nucleotide exchange activator homology F.37-89/Domain: SH3 homology <SH3B>
F.37-89/Domain: CDC25-type guanine nucleotide exchange activator homology <SOS>

```

Best Local Similarity 27.9%, Pred. No. 6,9e+08; Matches 63; Conservative 45; Mismatches 99; Indels 19; Gaps 6;

```
Oy      43  ERLVWNPQVFNELIPHDQGP-TVGSA-----EGLDVSAKDLAQLDTHDMSLFN  94
      1033 ERIVV---EKDPVPIPTKPPAPLTKGSLKLKPRVMDIDV---ELARQLTLREFKLYC 10855
Oy      95  SIHQVELIHYVLGPOHLDDVTYTNLEFERMRRELEQYVATLCLCPVGPRAOLLRKFI 154
Db      1086 KITPFACIACKWVGKSGKSGLSIESIDITQFIKASNOLDITFVNGVIMLRKADPKPKRVQDIIRFI 11445
Oy      155  KLAHLKEOKKLNSFFFAVMEGLSSASRLHATWERLPHKRYKLYSALERLIDPSMNIWR 214
Db      1146 QVADKCRQYNFSSMTAIIISALYSSPIHRLKKTWEYMAADALSNLKNMKNKLMSNRNFE 12055
Oy      215  YRLAKLT-SPPVIRFEMPIILKMDTFTHEGHTLVEN---LINPEK 256
Db      1206 YRDVLKFTGSEPCVPEFGVYISDLTEFYHNGRPDYLVNTRTOYVNAFK 1251
```

**RESULT 13**

S25716  
Ras guanine nucleotide exchange factor son-of-sevenless (sos) 1 - mouse  
N/Alternate names: Probable ras activator  
C/Species: Mus musculus (house mouse)  
C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 05-Nov-1999  
C/Accession: S25716; S21391  
R/Bowtell, D.; Fu, P.; Simon, M.; Senior, P.  
Proc. Natl. Acad. Sci. U.S.A. 89, 6511-6515, 1992  
A>Title: Identification of murine homologues of the Drosophila Son of sevenless gene:  
A/Accession number: S25714; MUID:92335328; PMID:1651150  
A/Molecule type: mRNA  
A/Residues: 1-1336 <ROW>

A/Cross-references: EMBL:Z11574; NID:g54134; PIDN:CAAT7662.1; PID:g54135  
C/Superfamily: CDC25-type guanine nucleotide exchange activator homology; plectstrln  
F:459-561/Domain: plectstrlin repeat homology <PLK>  
F:793-1036/Domain: CDC25-type guanine nucleotide exchange activator homology <SOS>

Query Match            11.0% Score 193; DB 2: Length 1336;  
Best Local Similarity     27.9%; Pred.No. 1,3e-07;  
Matches      60; Conservative     46; Mismatches       83; Indels          26; Gaps            9;

OY     65   PTV-----GSAEGLDLYSAK--DLAQQLTDHWSLFNSIHQVELHYVLGPQHRLDYT 115  
         |||           :|||       ::|||       :|||       |       :       ::  
DB     776   PTVEWHISRGHIETPOLLTLHPLEIRQLTTLESIDLRYAVQSSELVGWTKED--KEIN 834  
  
OY     116 TANERFRMRRENELQYWATELCCLPVPG--PRAOQLRKFKLAHLKEOKNLNFPAVM 173  
         :    :    :    :    :    :    :    :    :    :    :    :    :    :    :  
DB     835 SPNLAKMIRHTNTLTLMF--EKCIIVTENLEEVAVYSRIIEILQVFQELNNFGYLEVV 892  
  
OY     174 FGLSNSAISRAHWMLPKHKVRKYLSALELDPSNNHHYLRALAKL--SPVPYPM 230  
         ::::: |||||||::::: :    :    :    :    :    :    :    :    :    :  
DB     893 SAMMSFPVDRDHTEFDIOPSRKKTI---LEBAHELSBDH--YKKYLAKLNINSIPCVFF 947  
  
OY     231 PLLDKMTFHIEGNHSLV---ENLINFEKRMMMA 261  
         :    :    :    :    :    :    :    :    :    :    :    :    :    :  
DB     948 GIYLTNILKTPEGNPVELRRHGKLINFSSRRRAYA 982

**RESULT 14**

A37488  
Ras guanine nucleotide exchange factor son-of-sevenless (sos) 1 - human  
C/Species: Homo sapiens (man)  
C/Date: 03-Mar-1994 #sequence\_revision 18-Nov-1994 #text\_change 18-Mar-1997  
C/Accession: A37488  
R/Chardin, P.; Camonis, J.H.; Gale, N.W.; van Aeist, L.; Schlessinger, J.; Wigler, M.  
Science 260, 1338-1343, 1993  
A>Title: Human Sos1: a guanine nucleotide exchange factor for Ras that binds to GRB2.  
A/Reference number: A37488; MUID:9262494; PMID:8493579  
A/Accession: A37488  
A>Status: preliminary; not compared with conceptual translation  
A/Molecule type: mRNA





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 07:13:47 : Search time 25 seconds  
(Without alignments)  
393.787 Million cell updates/sec

Title: US-09-856-679-2

Perfect score: 1759  
Sequence: 1 MALALEDGWTGKGVLVKVN.....OLKVIDNORELSRLSRELEP 338

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 136899 seqs, 29126274 residues

Total number of hits satisfying chosen parameters: 136899

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending-Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/pae/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/pae/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/pae/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/pae/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/pae/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/pae/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/pae/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1392.5	79.2	819	1 PCT-US03-28227-3753	Sequence 3753, Ap
2	863.5	49.1	980	1 PCT-US03-28227-3136	Sequence 3136, Ap
3	637.5	36.2	456	6 US-10-435-696-97	Sequence 97, Appl
4	409	23.3	1499	7 US-60-480-890-348	Sequence 348, Appl
5	205.5	11.7	557	7 US-60-483-917-76	Sequence 76, Appl
6	193	11.0	1356	7 US-60-487-610-1523	Sequence 1523, Ap
7	193	11.0	1398	7 US-60-487-610-1524	Sequence 1524, Ap
8	188.5	10.7	1332	7 US-60-487-610-2054	Sequence 2054, Ap
9	182	10.3	699	6 US-10-240-145A-86	Sequence 86, Appl
10	182	10.3	715	6 US-10-240-145A-85	Sequence 85, Appl
11	147.5	8.4	821	6 US-10-296-115-1212	Sequence 112, Ap
12	108	6.1	814	7 US-60-487-610-1754	Sequence 1754, Ap
13	101.5	5.8	437	6 US-10-425-114A-46385	Sequence 46385, A
14	99	5.6	473	1 PCT-US03-22819-2	Sequence 2, Appl1
15	94.5	5.4	438	6 US-10-425-114A-40595	Sequence 40595, A
16	88	5.0	569	6 US-10-425-114A-64070	Sequence 64070, A
17	87.5	5.0	2816	6 US-10-240-145A-145	Sequence 145, Appl
18	87	4.9	271	5 US-09-897-516A-4785	Sequence 4785, Ap
19	87	4.9	2432	1 PCT-US03-28227-4702	Sequence 4702, Ap
20	87	4.9	2456	1 PCT-US03-28227-3997	Sequence 3997, Ap
21	87	4.9	2466	1 PCT-US03-28227-3996	Sequence 3996, Ap
22	87	4.9	2487	1 PCT-US03-28227-4001	Sequence 4001, Ap
23	87	4.9	2497	1 PCT-US03-28227-4000	Sequence 4000, Ap
24	87	4.9	2508	1 PCT-US03-28227-3999	Sequence 3999, Ap
25	87	4.9	2521	1 PCT-US03-28227-3998	Sequence 3998, Ap
26	85	4.8	533	6 US-10-425-114A-51223	Sequence 51223, A

27	84.5	4.8	382	5 US-09-920-923B-5	Sequence 5, Appl1
28	84.5	4.8	408	7 US-60-490-890-2256	Sequence 2256, Ap
29	84.5	4.8	760	7 US-60-490-890-2636	Sequence 2636, Ap
30	84	4.8	386	6 US-10-425-114A-59796	Sequence 59796, A
31	84	4.8	484	1 PCT-US03-10308A-29	Sequence 29, Appl
32	83.5	4.7	333	6 US-10-425-114A-51680	Sequence 51680, A
33	83	4.7	403	5 US-09-897-516A-7280	Sequence 7280, Ap
34	83	4.7	546	5 US-09-897-516A-4950	Sequence 4950, Ap
35	83	4.7	858	7 US-60-482-992-13	Sequence 13, Appl
36	83	4.7	926	7 US-60-480-890-81	Sequence 81, Appl
37	83	4.7	926	7 US-60-489-772-14	Sequence 14, Appl
38	83	4.7	926	7 US-60-498-106-2	Sequence 2, Appl1
39	83	4.7	1376	7 US-60-479-073-490	Sequence 490, Ap
40	82	4.7	481	7 US-60-487-610-2123	Sequence 2123, Ap
41	82	4.7	576	7 US-60-487-610-2124	Sequence 2124, Ap
42	81.5	4.6	311	6 US-10-425-114A-59795	Sequence 59795, A
43	81.5	4.6	530	6 US-10-425-114A-40263	Sequence 40263, A
44	81	4.6	603	6 US-10-415-656-32	Sequence 32, Appl
45	80.5	4.6	477	6 US-10-425-114A-65276	Sequence 65276, A

#### ALIGNMENTS

RESULT 1  
PCT-US03-28227-3753  
Sequence 3753, Application PC/TUS0328227

GENERAL INFORMATION:  
: APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;  
: APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;  
: APPLICANT: HARTJANOVIC, Mirjana M.; SHEN, Fan;  
: APPLICANT: MARGASHORE, Tolnette A.; SUCHOROSKI, Martin;  
: APPLICANT: ALFUS, Christina M.; PITTS, Steven J.;  
: APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;  
: APPLICANT: DELEGANE, Angelo M.; PANESAR, Iqbal S.;  
: APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;  
: APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;  
: APPLICANT: PANZER, Scott R.; WANG, Xinhuo;  
: APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;  
: APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;  
: APPLICANT: RIOUX, Pierre; SHEN, Edward J.;  
: APPLICANT: WU, Mingnam C.; STUVE, Laura L.;  
: APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;  
: APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;  
: APPLICANT: VIRT, Ursula A.; KIRTON, Edward;  
: APPLICANT: XU, Yuming; KWONG, Mary;  
: APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;  
: APPLICANT: MA, Yan; JACKSON, Jennifer L.;  
: APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;  
: APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.  
: TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS  
: FILE REFERENCE: PN-0100 PCT  
: CURRENT APPLICATION NUMBER: PCT/US03/28227  
: PRIOR FILING DATE: 2003-09-12  
: PRIOR APPLICATION NUMBER: US 60/410,260  
: PRIOR FILING DATE: 2002-09-12  
: PRIOR APPLICATION NUMBER: US 60/410,259  
: PRIOR FILING DATE: 2002-09-12  
: NUMBER OF SEQ ID NOS: 5444  
: SOFTWARE: PERL Program  
: SEQ ID NO 3753  
: LENGTH: 819  
: TYPE: PRT  
: ORGANISM: Homo sapiens  
: FEATURE:  
: NAME/KEY: misc-feature  
: OTHER INFORMATION: Incyte ID No: 932640.PT97P  
: PCT-US03-28227-3753

Query Match 79.2% Score 1392.5 DB 1 Length 819;  
Best Local Similarity 82.0% Pred. No. 1.6e-92;  
Matches 277; Conservative 0; Mismatches 0; Indels 61; Gaps 1;

```

QY 1 MALLAOEDGWTGQVLYVKNSSAGDAIGLOPDARGVATSLGNERLFVVPQVEHLEIHP 60
DB 543 MALLAOEDGWTGQVLYVKNSSAGDAIGLOPDARGVATSLGNERLFVVPQVEHLEIHP 602
QY 61 DOLGPTVGSAGEGLDVSADKGLDHDHMSLFNSIHQVELLHYVLGPOHLDDVTANLE 120
DB 603 DOLGPTVGSAGEGLDVSADKGLDHDHMSLFNSIHQVEHLEIHP 640
QY 121 RFRMRRENELOYWATELCLCPVGPRAQLLRKFIKLAHLKEOKNLNSFFAVMGISNSA 180
DB 641 -----LKEOKNLNSFFAVMGISNSA 661
QY 181 ISRLAHTWERLPHKVRKLYSALERLLDPSMNRHYRLAKLSPPVYIPFMPMLLKDMFTI 240
DB 662 ISRLAHTWERLPHKVRKLYSALERLLDPSMNRHYRLAKLSPPVYIPFMPMLLKDMFTI 721
QY 241 HEGNHTLVENLINFERRMARAARLHCRSHNPVPLSRVSHLHEDSOVARISTC 300
DB 722 HEGNHTLVENLINFERRMARAARLHCRSHNPVPLSRVSHLHEDSOVARISTC 781
QY 301 SEQSLSTRSPASTWAVYVQOLKVIDNORELSTRSRELEP 338
DB 782 SEQSLSTRSPASTWAVYVQOLKVIDNORELSTRSRELEP 819

```

RESULT 2  
PCT-US03-28227-3136

; Sequence 3136, Application PC/TUS0328227

; GENERAL INFORMATION:

```

; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;
; APPLICANT: ALTUS, Christine M.; PITTS, Steven J.;
; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
; APPLICANT: DELEGEANE, Angelo M.; PANESAR, Iqbal S.;
; APPLICANT: BANYILLE, Steven C.; REDDY, Thirupathi P.;
; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;
; APPLICANT: PANZER, Scott R.; WANG, Xinhao;
; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
; APPLICANT: PERALTA, Careyua H.; ANDERSON, Scott E.;
; APPLICANT: RIOUX, Pierre; SHEN, Edward J.;
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITTE, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuming; KWONG, Mary;
; APPLICANT: POLICKY, Jennifer L.; HUBWITZ, Bonnie L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/28227
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO 3136
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 928293.PT40P
; NAME/KEY: unsure
; LOCATION: (1) ... (980)
; OTHER INFORMATION: unknown or other
; PCT-US03-28227-3136

```

Query Match 49.1%; Score 863.5; DB 1; Length 980;  
Best Local Similarity 50.0%; Pred. No. 1,3e-54;  
Matches 169; Conservative 63; Mismatches 81; Indels 25; Gaps 3;

```

QY 1 MALLAOEDGWTGQVLYVKNSSAGDAIGLOPDARGVATSLGNERLFVVPQVEHLEIHP 60
DB 665 ISAVADKLSGEGLLIVKSSGGEKVYLPNDVSVFTTLTNGRLFACPREFDSLTPLP 724
QY 61 DOLGPTVGSAGEGLDVSADKGLDHDHMSLFNSIHQVELLHYVLGPOHLDDVTANLE 120
DB 725 DOLGPTVGSAGEGLDVSADKGLDHDHMSLFNSIHQVELLHYVLGPOHLDDVTANLE 783
QY 121 RFRMRRENELOYWATELCLCPVGPRAQLLRKFIKLAHLKEOKNLNSFFAVMGISNSA 180
DB 784 LFRRENELOFWVWTEICISQSLSRVQLKFKFIKLAHCKEYKNLNSFFAVMGISNSA 843
QY 181 ISRLAHTWERLPHKVRKLYSALERLLDPSMNRHYRLAKLSPPVYIPFMPMLLKDMFTI 240
DB 844 VSRLALTWKELRSKKEFAEFESLMDPSRNRHAYRLAVAKLEPLIPFMPMLLKDMFTI 903
QY 241 HEGNHTLVENLINFERRMARAARLHCRSHNPVPLSRVSHLHEDSOVARISTC 300
DB 904 HEGNHTLVENLINFERRMARAARLHCRSHNPVPLSRVSHLHEDSOVARISTC 953
QY 301 SEQSLSTRSPASTWAVYVQOLKVIDNORELSTRSRELEP 338
DB 954 -----SVRQOLNVIDNOFTLSQSHRELEP 977

```

RESULT 3  
US-10-435-696-97

; Sequence 97, Application US/10435696

; GENERAL INFORMATION:

```

; APPLICANT: Wirtz, Ralph
; APPLICANT: Munnes, Marc
; APPLICANT: Kallabis, Harald
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE PREDICTION, DIAGNOSIS, PROGNOSIS
; FILE REFERENCE: Lea 36 108
; CURRENT APPLICATION NUMBER: US/10/435,696
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: EP03003112.4
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: EP02010291.9
; PRIOR FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 314
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-435-696-97

```

Query Match 36.2%; Score 637.5; DB 6; Length 456;  
Best Local Similarity 42.0%; Pred. No. 9.1e-39;  
Matches 136; Conservative 60; Mismatches 99; Indels 29; Gaps 6;

```

QY 15 VLVKNSAGDAIGLOPDARGVATSLGNERLFVVPQVEHLEIHPDOLGPTVGSAGEGLD 74
DB 157 ILVAVSSGSEKVLDPTEDCVFTALGINSHPACTRDSYEALVPLPEELOVSPGOTE-IH 215
QY 75 LVSAKDAGOLTDHMSLFNSIHQVELLHYVL-GPOHLDDVTANLEFRMRRENELOYW 133
DB 216 RVEPDEVANHLTAFAHMELEFRCHHELEFDVYHNGRGRE--TANLELLOQCSVTHV 273
QY 134 ATELCVPGPRAQLLRKFIKLAHLKEOKNLNSFFAVMGISNSAISRLAHTWERLPH 193
DB 274 ATEVLCEAPGRRAQLLRKFIKLAHLKEOKNLNSFFAVMGISNSAISRLAHTWERLPH 333
QY 194 KVRKLYSALERLLDPSMNRHYRLAKLSPPVYIPFMPMLLKDMFTIHGNTLVENLIN 253
DB 334 KFRNLEFRFENLTDCRNHKSYSREVISKMKPPVPLILKDLTFLEHSGKTLVDGLVN 393

```



RESULT 7  
US-60-487-610-1524  
; Sequence 1524, Application US/60487610  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: HUANG, Hongjin  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,  
; FILE REFERENCE: C1001469  
; CURRENT APPLICATION NUMBER: US/60/487,610  
; CURRENT FILING DATE: 2003-07-17  
; NUMBER OF SEQ ID NOS: 97101  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1524  
; LENGTH: 1398  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-487-610-1524

Query Match 11.0%; Score 193; DB 7; Length 1398;  
Best Local Similarity 27.9%; Pred. No. 1.5e-06;  
Matches 60; Conservative 46; Mismatches 83; Indels 26; Gaps 9;

QY 65 PTV-----GSAEGLDVSAR--DLAQLTDHMSLFSNHQVELIHVGLGPQHLRDVT 115  
DB 839 PTVWHISRPGHIEFDLTLHPRIARQLLESDLYRAVQPSLVSQVMTKED--KEIN 897  
QY 116 TANLEFRMRRELOYWATELCCLPVG--PRAQLLKFKLAHKEOKNLSFFAVM 173  
DB 898 SPNLKMRHRTNTLTMF--EKCIVETENLEERAVASRIIEILOVFOELNFGVLEVV 955  
QY 174 FGLSATSISRLAHWTWERLPHKVKRLYSALERLLDPSNMHRVRYRLAKL---SPVVIDFM 230  
DB 956 SAMNSPYRLDHTFEQJPSROKTI---LEENHEISEH--YKTLAKLRISINPCVPEF 1010  
QY 231 PLLKDMTFIHGNTLV---ENLINEKRMMA 261  
DB 1011 GYLTNIIKTEGNEPEVLKRGKELINFSKRRKVA 1045

RESULT 8  
US-60-487-610-2054  
; Sequence 2054, Application US/60487610  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: HUANG, Hongjin  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,  
; FILE REFERENCE: C1001469  
; CURRENT APPLICATION NUMBER: US/60/487,610  
; CURRENT FILING DATE: 2003-07-17  
; NUMBER OF SEQ ID NOS: 97101  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2054  
; LENGTH: 1332  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-487-610-2054

Query Match 10.7%; Score 188.5; DB 7; Length 1332;  
Best Local Similarity 24.8%; Pred. No. 3.1e-06;  
Matches 60; Conservative 53; Mismatches 102; Indels 27; Gaps 9;

QY 29 QPDAAGVATSLGLNERLFVNPQVEHLLIPPDQLGPTVGSAGGLDVSAR--DLAQLT 86  
DB 740 QAQANGVSHNT-----TFESPPELWHISKP-----GQFETFDLTLHPRIARQLT 787  
QY 87 DHMSLFSNHQVELIHVGLGPQHLRDVTANLEFRMRRELOYWATELCCLPVG-- 144

DB 788 LIESDLYKVPQPSLVSQVMTKED--KEINSPNLKMRHRTNTLTMF--EKCIVAEENFE 844  
QY 145 PRAQLLKFKLAHKEOKNLSFFAVMGLSATSISRLAHWTWERLPHKVKRLYSALER 204  
DB 845 ERVAVLSRIIELOVFOELNFGVLEIVSAVNSVSYRLDHTFEALQERRKTI---LDE 901  
QY 205 LIDPSWNH--RYRLALAKLSPVPIPEMPLDKMTFIEGNTLV---ENLINEKRM 259  
DB 902 AVELSDQHFKKRYLKLKSNPCVPEFGIYLTNIIKTEGNNDELKKGKGLINFSKRRK 961  
QY 260 MA 261  
DB 962 VA 963

RESULT 9  
US-10-240-145A-86  
; Sequence 86, Application US/10240145A  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-048  
; CURRENT APPLICATION NUMBER: US/10/240,145A  
; CURRENT FILING DATE: 2002-09-27  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: 09/668,680  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 09/695,618  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 09/728,711  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED  
; PRIOR FILING DATE: 2000-03-14  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: Custom  
; SEQ ID NO 86  
; LENGTH: 699  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-240-145A-86

Query Match 10.3%; Score 182; DB 6; Length 699;  
Best Local Similarity 23.7%; Pred. No. 5e-06;  
Matches 75; Conservative 52; Mismatches 118; Indels 72; Gaps 11;

QY 51 QVEHLLIPHPDQLGPTVGSAGGLDVSARKDLAQLTDHMSLFSNHQVELIHVGLGPQ 110  
DB 236 EEEGLMPOGROL-----LDF-SVDEVAEQLTIDELFESKVRUYECIGSWSORD 285  
QY 111 LRDVTANN--DERMRRENELOYWATELCCLP--VPGP-RAQLRKFKLAHKEOKNL 166  
DB 286 RPGAAGASPYRAVAOENITVGCYLSVGLAPBLAARORAROREKIRIAORRELRNF 345  
QY 167 NSFPAVVGSLNSAISRIAHWTWERLPHKVKRLYSALERLLDPSNMHRVRYRLAT----- 219  
DB 346 SSLRAIISALQSNITVRLKRWGAVSREPLSTFKLSQISDENNHLSREILFQEBATE 405  
QY 220 -----AKLSPPVPEPMLLKDMTFIEGNTLVF--NLINFEKMR---MA 261  
DB 406 GSOEDMTPGSLPSKPPGPVPVYGLTLDVMDTLAPMLLEGLDINFEKREWEILA 465  
QY 262 RAAMLIHCRHNHNPVPLSRVSHLHEDSOVARISTCSROSISTSPASTMAYVQQLK 321  
DB 466 RIQQLORRCOSYTLSPHPPI---LAALHAONQLT-----EBOSY----- 501  
QY 322 VIDMQRELSRLELEP 338  
DB 502 -----RLSRVIEP 509

```
RESULT 10
; Sequence 85, Application US/10240145A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-048
; CURRENT APPLICATION NUMBER: US/10/240,145A
; CURRENT FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/668,680
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,618
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 09/728,711
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: Custom
; SEQ ID NO 85
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-240-145A-85
```

```
Query Match 10.3%; Score 182; DB 6; Length 715;
Best Local Similarity 23.7%; Pred. No. 5.1e-06;
Matches 75; Conservative 52; Mismatches 118; Indels 72; Gaps 11;
```

```
QY 51 QEVHELPHDQGPVGSNEGDLVSAKDLAGQLDHDMSLNSHNOVELHYVAGPQH 110
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 236 EEEEGMPPQPOQ-----LDF-SVEVAGQLTLDLEFSKYLVECLGSVMSQRD 285
QY 111 LRDVTAN--LEFRMRFNELQYVATTELCP-VGCP-RAQLRKFIKLAHLKBOKNL 166
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 286 RPAAGASPVYATVQAFNTVTCVGLSGVAGGLAPQAOALEKIRIAQOCRELRF 345
QY 167 NSFFAVMFGLSNSAISRLAHTWBLPHKVKLYSALERLDPMSNHRVYRLAL----- 219
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 346 SSLRALISALQSNPIYRLKRSKMGAVSREPLSTFRKLSQIRSDSNHLSSEILFOEATE 405
QY 220 -----AKSPPIYIPMPLLDKMTFIEGNTLVE-NLINFEMR-----MMA 261
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 406 GSGEDNTPGSLRKPDPGPVPLGTFTLDVMDLDPMLDGLDINFERRKWEIILA 465
QY 262 RAARMLHGRSHNPVPLSRVSHLHEDSOVARISTCSQSLSRSPASTWAVYQOLK 321
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 466 RIQQLORCOSYLSLSPHPF-----LAALHAONOLT-----EESY----- 501
QY 322 VIDNORELSRLSRELPEP 338
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 502 -----RLSRVIEP 509
```

```
RESULT 11
US-10-296-115-1212
; Sequence 1212, Application US/10296115
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/486,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1212
```

```
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-1212
```

```
Query Match 8.4%; Score 147.5; DB 6; Length 821;
Best Local Similarity 21.7%; Pred. No. 0.0017;
Matches 66; Conservative 49; Mismatches 118; Indels 71; Gaps 10;
```

```
QY 78 AKDL-AGQLTDHMSLFPNSHNOVELHYVUG-----PQHLRDYTTANLERPMRRP 126
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 284 SEDLVAEQLTYMDQLEFKV-----VPHHCIGCIMSRRDKENKHNLAFTJATIS----QP 335
QY 127 NELQYVATTELCPVPG-----PRAOLRKFIKLAHLKBOKNLNSFFAVMFGLSNS 179
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 336 NTL-----TKCVSTLGGKELTQOAKITEKWINAHECRLLKNFSSLRATVALQSN 350
QY 180 AISRLAHTWBLPHKVKLYSALERLDPMSNHRVYRLALAK----- 221
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 391 SIYRLAKTMAAVPRDRLMFEELSDIFSDHNNHLSRELLMKEGTGSFANLDSSVKNOK 450
QY 222 -----LSPPVLPFMPPLLDKMTFIEGNTLVE-NLINFEMR-----MMA 263
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 451 RTQRRLOQDKMGVMQGYVPLGTFTLDLTMLTALQDYIEGLINFEKRRREFVYAOI 510
QY 264 ARMLHGRSHNPV-----LSPLRSVSHLHEDSOVARISTCSQSLSRSPASTWAVYQO 319
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 511 KILQSNCSYCMTPDQKFIOMFOROQLTEESIALCEIAMAADSTSPKPKMSVKKR 570
QY 320 LKVI 323
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 571 LNL 574
```

```
RESULT 12
US-60-487-610-1754
; Sequence 1754, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1754
; LENGTH: 814
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-1754
```

```
Query Match 6.1%; Score 108; DB 7; Length 814;
Best Local Similarity 21.9%; Pred. No. 1.1;
Matches 79; Conservative 50; Mismatches 120; Indels 112; Gaps 16;
```

```
QY 31 DANGVATSLGLNE--RLEVVNPQEVHELPHDQGLPVGSAAGDLVSAKDLAQLTDH 88
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 342 DRGQVITMQALSEDRRLMEAMDGRPYVNSNKSQSESTA-----QDSI 388
QY 89 DMSLFNS-INQVELIHVILGPOHLRDYTTANLERPMRFELOYWA-----T 135
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 389 GFSSIRKCHAVE-----TRGINEOGLYRIGVNSRVOKLLSLVMDPITASSETP 438
QY 136 ELG-----LCVPVGP--RAOLRKFIKLAHLKBOKNLNSFFAVMFGLSN 178
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 439 DICAWEIKITTSALKTYLMLRGPRLMYPQFSFIK-AKLENQ----- 483
QY 179 SAISRLAHTWBLPHKVKLYSAL-ERLDPMSNHRVYRLALAKLSPPVLPFMPPLIK-- 235
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 484 SRVSEIHSIVHRRLPEKRRQMLQLMNLHANVANNHKKONLMTVANLG---VFGPPTLLRPQ 540
```





**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 07:11:44 : Search time 29 Seconds  
(without alignments)  
493.140 Million cell updates/sec

Title: US-09-856-679-2  
Perfect score: 1759  
Sequence: 1 MAALAEDEGWTGQVLKVN.....QLKVIDNORELSRLSELP 338

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/laa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/laa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/laa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/laa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/laa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/laa/Backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	213	12.1	557	US-09-568-816A-2	Sequence 2, App11
2	208.5	11.9	911	US-09-356-952-6	Sequence 6, App11
3	206	11.7	1244	US-09-356-952-7	Sequence 7, App11
4	202.5	11.5	402	US-08-290-731C-14	Sequence 14, App11
5	199	11.3	489	US-08-318-831-4	Sequence 4, App11
6	199	11.3	666	US-08-318-831-3	Sequence 3, App11
7	199	11.3	814	US-08-318-831-2	Sequence 2, App11
8	194	11.0	364	US-08-318-831-6	Sequence 6, App11
9	193	11.0	423	US-08-290-731C-10	Sequence 10, App11
10	193	11.0	684	US-09-765-298A-18	Sequence 18, App11
11	193	11.0	1319	US-08-290-731C-2	Sequence 2, App11
12	193	11.0	1333	US-09-356-952-2	Sequence 2, App11
13	193	11.0	1336	US-08-290-731C-6	Sequence 6, App11
14	189.5	10.8	652	US-08-318-831-8	Sequence 8, App11
15	181	10.3	418	US-08-290-731C-13	Sequence 13, App11
16	181	10.3	1048	US-09-356-952-5	Sequence 5, App11
17	178	10.1	426	US-08-290-731C-12	Sequence 12, App11
18	178	10.1	430	US-08-290-731C-5	Sequence 5, App11
19	178	10.1	1572	US-08-280-731C-9	Sequence 9, App11
20	178	10.1	1589	US-09-356-952-4	Sequence 4, App11
21	178	10.1	1596	US-09-356-952-3	Sequence 3, App11
22	177.5	10.1	423	US-08-290-731C-11	Sequence 11, App11
23	177.5	10.1	1297	US-08-290-731C-4	Sequence 4, App11
24	164.5	9.4	362	US-08-280-731C-15	Sequence 15, App11
25	153.5	8.7	852	US-08-408-519-5	Sequence 5, App11
26	153.5	8.7	852	PCT-US95-03552-5	Sequence 5, App11
27	143	8.1	768	US-08-408-519-2	Sequence 2, App11

28	143	8.1	768	5	PCT-US95-03552-2	Sequence 2, App11
29	127.5	7.2	688	4	US-09-367-206-20	Sequence 20, App11
30	127.5	7.2	703	4	US-09-367-206-5	Sequence 5, App11
31	121.5	6.9	501	4	US-09-367-206-3	Sequence 3, App11
32	103.5	5.9	1504	4	US-09-364-206-2	Sequence 2, App11
33	96	5.5	337	4	US-09-252-991A-25798	Sequence 25798, A
34	95.5	5.4	452	4	US-09-252-991A-26867	Sequence 26867, A
35	93	5.3	462	3	US-09-055-113-4	Sequence 4, App11
36	91	5.2	486	3	US-08-348-518C-2	Sequence 2, App11
37	91	5.2	515	3	US-08-930-966A-12	Sequence 12, App11
38	88.5	5.0	473	1	US-08-103-739B-2	Sequence 2, App11
39	88.5	5.0	473	2	US-08-474-404-2	Sequence 2, App11
40	88.5	5.0	473	2	US-08-485-845-2	Sequence 2, App11
41	88.5	5.0	473	2	US-08-482-714-2	Sequence 2, App11
42	88.5	5.0	473	3	US-09-211-416-2	Sequence 2, App11
43	88.5	5.0	473	3	US-09-059-858-2	Sequence 2, App11
44	88	5.0	639	4	US-09-252-991A-25089	Sequence 25089, A
45	84.5	4.8	382	3	US-08-660-645A-7	Sequence 7, App11

ALIGNMENTS

RESULT 1  
US-09-568-816A-2  
: Sequence 2, Application US/09568816A  
: Patent No. 6440699  
: GENERAL INFORMATION:  
: APPLICANT: Tavitligian, Sean V.  
: APPLICANT: Swedlund, Brad  
: APPLICANT: Simard, Jacques  
: APPLICANT: Rommens, Johanna M.  
: APPLICANT: Myriad Genetics, Inc.  
: APPLICANT: Hospital for Sick Children  
: TITLE OR INVENTION: CA7 CG04 Gene  
: FILE REFERENCE: 2318-237-11  
: CURRENT APPLICATION NUMBER: US/09/568, 816A  
: CURRENT FILING DATE: 2000-05-11  
: PRIOR APPLICATION NUMBER: US 60/134, 209  
: PRIOR FILING DATE: 1999-05-14  
: NUMBER OF SEQ ID NOS: 21  
: SOFTWARE: Patentln Ver. 2.0  
: SEQ ID NO 2  
: LENGTH: 557  
: TYPE: PRT  
: ORGANISM: Homo sapiens  
: US-09-568-816A-2

Query Match 12.1%; Score 213; DB 4; Length 557;

Best Local Similarity 27.1%; Pred. No. 4.2e-14; Matches 70; Conservative 45; Mismatches 113; Indels 30; Gaps 7;

QY	128	ELOYWATELCLCPVGPRAQLLRKFIKLAHLKEQNLSFPAVMGLSNSAISLAHT	187
DB	48	VTFEEAQTLLMDPVFEKAIOPDELSGCMNKKERYSSAP-----NAVAFTRREN	98
QY	99	HSEFWVREILHAOTLKIRAEVLSHYITAKKLYELNNLHMLMVSLDSAPITFLTKT	158
DB	188	WERLPKHKVRLKSALERLDPSNMNRHVYRLAKLSL-PVYIPEFPLILKDMTFHECNIHT	246
QY	159	WALLSKRDKTTFEEKLEYVMSKEDNKKRLROYISSLKMPCCIPYIGIYLSDTYIDSAVPS	218
DB	247	LVENLINFKKMMMAARAARMLHHGRS--HNPPVLSPP---LRSVSHNE-----DSQVA	295
QY	219	TGSILENORSNLMMNLIIRISDLOOSEYDILPHLPHYQKLYNSVQYIEELQKVEDNY	278
QY	296	RISTCSQSLSLT-RSPAS	312
DB	279	KUSLKIEPTSTPRSAAS	296











TITLE OF INVENTION: PREPARATION AND UTILIZATION  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (EPO PatentIn)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/318, 831  
FILING DATE: 19 October 1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR92/04827  
FILING DATE: 21-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92033-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 652 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-318-831-8

Query Match 10.8%; Score 189.5; DB 1; Length 652;  
Best Local Similarity 24.8%; Pred. No. 1.8e-11;  
Matches 60; Conservative 53; Mismatches 102; Indels 27; Gaps 9;

QY 29 QDARQVATSLGNEELFVYNQVEHLIRHPDQLPTVGSAGLDLVSAK--DLGQLT 86  
DB 97 QAAQANSHNI-----TFESPPIPIEMHISKP-----GQETFDLMLHPLEIAROLT 144  
QY 87 HDWMSLFNSIHQVELIHYVGLPOHLRDVTTANLERMRERNELOVWVATELCLCPVGC-- 144  
DB 145 LLESIDYKRVQPSSELVGSVWTKED-KEINSPNLKIRHTNLTLMF--EKCIIVEENEE 201  
QY 145 PRAQLRKFKIKLAHLKEQKNLNSFPAVMEGLSNSAISRLATWERTLPHKVRKLYSALER 204  
DB 202 ERVAVALSRILEILOVRDLNFNGLVEIYSAVNSVYRLDHTFEALQEKRRKI--LDE 258  
QY 205 LLDPSNNH-RVYRLALAKLSPYVIRPMPILLKDMTIEHGNHTLV-----ENLINFEMRM 259  
DB 259 AVELSDHFKKYLVLKLSINPCVPFPGIYLTNLTKEBGNDFLKKKGKDLINFSKRRK 318  
QY 260 MA 261  
DB 319 VA 320

RESULT 15  
US-08-290-731C-13  
Sequence 13. Application US/08290731C  
Patent No. 5843646  
GENERAL INFORMATION:  
APPLICANT: BOWTELL, David Douglas Lawrence  
TITLE OF INVENTION: DNA MOLECULES ENCODING MURINE  
TITLE OF INVENTION: SON OF SEVENTEEN (MSOS) GENE,  
TITLE OF INVENTION: AND MSOS POLYPEPTIDES  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
STREET: 2100 PENNSYLVANIA AVENUE, N.W.

CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/290, 731C  
FILING DATE: 17-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU93/00068  
FILING DATE: 17-FEB-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PL0921/92  
FILING DATE: 17-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: KIT, Gordon  
REGISTRATION NUMBER: 30,764  
REFERENCE/DOCKET NUMBER: Q-36066  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 293-7060  
TELEFAX: (202) 293-7860  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 418 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-290-731C-13

Query Match 10.3%; Score 181; DB 2; Length 418;  
Best Local Similarity 23.2%; Pred. No. 7.4e-11;  
Matches 73; Conservative 57; Mismatches 144; Indels 40; Gaps 8;

QY 14 QVLVKNSSAGDAIGLDPDARGVATSLGNEELFVYNQVEHLIRHPDQ-----LGPT 66  
DB 107 QVAIKENINGSVELLE-----VNQKFLGNIOETATPMKTLDDOICODHYSCTL 156  
QY 67 VGSAGGLDVSADKDLGOLTDHDWMSLFNSIHQVELIHYVGLPOHLRDV--TTANLERMR 125  
DB 157 YSTTESILAVDPVLFATQTLTLEHEIYCELTITDCLQIKKNTSYSGASPOLNEFISF 216  
QY 126 FNELOVWVATELCLCPVGPRAQLRKFKIKLAHLKEQKNLNSFPAVMEGLSNSAISRLA 185  
DB 217 ANKLTFNISYSYVKKADSKRAKLGHFIFIAEYCKRFNFSMTDIIISALYSPIYRL 276  
QY 186 HTWERLPHKVRKLYSALERLDDPSMNRVYRLALAKL-SPYVIRPMPILLKDMTIEHGN 244  
DB 277 KTWQAAVLPQTRDLOSINKLMDPKNFINRNLKSLHSPCVPFPGVYLSDLTFDSDGN 336  
QY 245 -----HTL-----VENLINFEK--MRMARAARMLHHCNSHPVP-----LSPLRS 283  
DB 337 PDYLVLEHGLKGVHDEKKTINENKRSRLVDIIOELITTFKKTHTDFTKDRIVICISNSLE 386  
QY 284 RVSHLEDSDQARI 297  
DB 397 NIPHEKQYQSLI 410

Search completed: October 7, 2003, 07:16:48  
Job time: 31 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 06:56:52 : Search time 86 Seconds  
(without alignments)  
623.832 Million cell updates/sec

Title: US-09-856-679-2

Perfect score: 1759  
Sequence: 1 MAALAQEDGKTVGLVKNV.....QLKVIDNRELSTRLELP 338

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*
- 24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1759	100.0	338	21	AAV99651 Human GTPase assoc
2	1748	99.4	881	21	AAV70965 Human Ras signalling
3	1664	94.6	884	21	AAV70964 Rat Ras signalling
4	863.5	49.1	849	21	AAV70968 Human Ras signalling
5	786.5	44.7	291	21	AAV70967 Rat Ras signalling
6	696	39.6	957	22	AB59888 Drosophila melanog
7	637.5	36.2	456	22	AAW8706 Human Protein SEQ
8	637.5	36.2	456	22	AAW8706 Human polypeptide
9	419.5	23.8	246	23	AAE22105 Human 48921 guanin

10	409	23.3	1499	21	AA07792 A murine guanine n
11	405.5	23.1	261	22	AAE3102 Human phospholipase
12	405.5	23.1	261	22	AAU99911 Human guanin nucle
13	402.5	22.9	1551	22	AB564459 Drosophila melanog
14	390.5	22.2	1651	22	AB566725 Human novel polype
15	390.5	22.2	1675	21	AA842658 Human ORFX ORF2422
16	283	16.1	1077	16	AA891597 C3g protein. Homo
17	261	14.8	1380	22	AB595638 Drosophila melanog
18	246.5	14.0	227	22	AAU99920 Human 47476 consen
19	237	13.5	472	22	AA578823 Human guanin-nucl
20	237	13.5	472	23	AB804984 Human new ras guan
21	234	13.3	665	22	AB561569 Drosophila melanog
22	234	13.3	702	22	AB561513 Human liver peptid
23	229	13.0	59	22	AB55062 Human 47435 enco
24	229	13.0	59	22	AB839929 protein #6475 enco
25	229	13.0	59	22	AB824476 Human bone marrow
26	229	13.0	59	22	AAW60674 Human brain expres
27	229	13.0	59	22	AAW73346 Peptide #6357 enco
28	229	13.0	59	22	AAW19923 Peptide #7585 enco
29	229	13.0	59	22	AAW3548 Human peptide enco
30	229	13.0	59	23	AB543199 Human GTP-binding
31	228.5	13.0	466	22	AB568566 Human Ras releasin
32	227	12.9	472	23	AAE22102 Drosophila melanog
33	225	12.8	612	22	AB871554 Ras-like related h
34	221	12.6	591	23	AA026358 Human polypeptide
35	214	12.2	1237	22	AAW39825 Human GPF2 protein
36	214	12.2	1237	23	AAU11805 Human polypeptide
37	214	12.2	1257	22	AAW41611 Novel human diagno
38	214	12.2	1268	22	AB515297 Human polypeptide
39	213	12.1	361	22	AAW40386 Human prostate can
40	213	12.1	557	22	AA848789 Novel human protei
41	213	12.1	557	23	AB897502 Human protein sequ
42	211	12.0	709	22	AB562626 Drosophila melanog
43	208.5	11.9	274	22	AB561034 Amino acid sequenc
44	208.5	11.9	911	21	AA568824 Breast cancer asso
45	207.5	11.8	782	24	ABR47580

#### ALIGNMENTS

RESULT 1	AAV99651 standard; Protein: 338 AA.
ID	AAV99651
XX	03-NOV-2000 (first entry)
AC	Human GTPase associated protein-2.
XX	
AC	AAV99651;
XX	
DT	03-NOV-2000 (first entry)
XX	
DE	Human GTPase associated protein-2.
XX	
KW	Guanine nucleotide binding protein; GTP-binding protein; G-protein;
KW	GTPase; GTPase associated protein; GTPAP; cell proliferation;
KW	autoimmune; inflammatory; immune system disorder; cancer; AIDS;
KW	acquired immune deficiency syndrome; asthma; atherosclerosis;
KW	arthritis; systemic lupus erythematosus; psoriasis; human.
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FT	Modified-site 21
FT	/note= "Potential phosphorylation site"
FT	Modified-site 77
FT	/note= "Potential phosphorylation site"
FT	Modified-site 86
FT	/note= "Potential phosphorylation site"
FT	Modified-site 131
FT	/note= "Potential phosphorylation site"
FT	Modified-site 200
FT	/note= "Potential phosphorylation site"
FT	Modified-site 244
FT	/note= "Potential glycosylation site"
FT	Modified-site 246

```

FT Modified-site /note="Potential phosphorylation site"
FT 299 /note="Potential phosphorylation site"
FT Modified-site 306
FT /note="Potential phosphorylation site"
XX
XX MO200031263-A2.
XX
XX
XX 02-JUN-2000.
XX
XX 23-NOV-1999; 99WO-US28013.
XX
XX 23-NOV-1998; 98US-0109592.
XX 04-FEB-1999; 99US-0118610.
XX 06-APR-1999; 99US-0127990.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Hillman JL, Tang YT, Bandman O, Lal P, Yue H, Lu DM, Baughn MR;
XX Yang J, Azimzai Y;
XX
XX WPI: 2000-400073/34.
XX N-PSDB; AAA49172.
XX
XX Human GTPase associated proteins, polynucleotides, and antibodies,
XX useful for diagnosing, preventing and treating various diseases such as
XX atherosclerosis, cancer, acquired immune deficiency syndrome (AIDS),
XX asthma, and autoimmune diseases -
XX
XX Claim 1: Page 83-84; 144pp; English.
XX
XX Human cDNA libraries from various tissues were screened for GTPase
XX associated proteins (GTPAP). The present sequence is human
XX GTPAP-2 protein. This sequence was derived from a cDNA library of the
XX brain meningioma tissue from a 35 year old female. This protein is
XX expressed in reproductive, nervous and cardiovascular tissue. The
XX GTPAP proteins may be used to define agonists and antagonists of GTPAP
XX activity and to generate antibodies to GTPAP. This means the GTPAP
XX proteins may be useful for treatment or prevention of diseases
XX associated with GTPAP such as cell proliferation disorders, autoimmune
XX disorders, inflammatory disorders, immune system disorders, cancer,
XX AIDS, asthma, atherosclerosis, arthritis, systemic lupus erythematosus
XX and psoriasis.
XX
XX Sequence 338 AA:
SQ
Query Match 100.0%; Score 1759; DB 21; Length 338;
Best Local Similarity 100.0%; Pred. No. 1,8e-169;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 2
ID AAY70965
XX AAY70965 standard; Protein; 881 AA.
XX
XX AAY70965;
XX
XX 09-AUG-2000 (first entry)
XX
XX Human Ras signalling pathway associated protein CAMP-GEFI.
XX
XX Human; Ras signalling pathway; CAMP-GEFI; cyclic adenosine monophosphate;
XX GEF; guanine nucleotide exchange factor; Rap1A; diagnosis; treatment;
XX CAMP-GEF-associated disorder; drug; transgenic animal model;
XX Ras-associated cancer; protein therapy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Binding-site 231..300
XX /label=CAMP-binding_domain
XX
XX Region 278..282
XX /note="conserved motif in the CAMP binding pocket;
XX amino acid at position 280 determines binding
XX specificity for CAMP/CGMP"
XX
XX Region 616..639
XX /label=SCR1
XX /note="Structurally conserved region which is
XX highly homologous to Ras-superfamily GEFs"
XX
XX Region 689..731
XX /label=SCR2
XX /note="Structurally conserved region which is
XX highly homologous to Ras-superfamily GEFs"
XX
XX Region 768..789
XX /label=SCR3
XX /note="Structurally conserved region which is
XX highly homologous to Ras-superfamily GEFs"
XX
XX MO200024768-A2.
XX
XX 04-MAY-2000.
XX
XX 22-OCT-1999; 99WO-US24826.
XX
XX 23-OCT-1998; 98US-0105507.
XX 16-NOV-1998; 98US-0108685.
XX
XX (MAST ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Kawasaki H, Graybiel A, Housman D;
XX WPI: 2000-350690/30.
XX N-PSDB; AAD00315.
XX
XX Isolated nucleic acid comprises nucleotide sequence encoding protein
XX selected from normal or mutant CalDAG-guanine nucleotide exchange
XX factor
XX
XX Claim 64; Page 113-115; 128pp; English.
XX
XX The present sequence is a human cyclic adenosine 3', 5'
XX monophosphate-guanine nucleotide exchange factor (CAMP-GEFI) which
XX has CAMP binding domain and Ras superfamily GEF
XX domains. It has substrate specificity for Rap1A and is differentially
XX distributed in brain and various other tissues. It selectively activates
XX the Ras superfamily small G protein and Rap1A, functions as a negative
XX regulator of Ras and directly couples the CAMP signal transduction
XX system to Ras superfamily cascades. The present sequence
XX is used for diagnosis and treatment (by protein therapy) of
XX CAMP-GEF-associated disorders. It is also useful for developing drugs
XX and producing cell-lines or transgenic animal models for Ras-associated
XX cancers.

```

[illegible]

XX	RESULT 3
XX	AAV70964
XX	AAV70964 standard; Protein: 884 AA.
XX	AAV70964;
XX	09-AUG-2000 (first entry)
XX	Rat Ras signalling pathway associated protein CAMP-GEFI.
XX	
XX	Rat: Ras signalling pathway: CAMP-GEFI; cyclic adenosine monophosphate;
XX	GEP; guanine nucleotide exchange factor; RAP1A; diagnosis: treatment;
XX	CAMP-GEF-associated disorder; drug; transgenic animal model;
XX	Ras-associated cancer; protein therapy.
XX	Rattus norvegicus.
XX	
XX	Key
XX	Location/Qualifiers
XX	231..300
XX	/label= CAMP-binding_domain
XX	278..282
XX	/note= "conserved motif in the CAMP binding pocket;
XX	amino acid at position 280 determines binding
XX	specificity for CAMP/CGMP"
XX	619..642
XX	/label= SCR1
XX	/note= "Structurally conserved region which is
XX	highly homologous to Ras-superfamily GEFs"
XX	692..734
XX	/label= SCR2
XX	/note= "Structurally conserved region which is
XX	highly homologous to Ras-superfamily GEFs"
XX	771..792
XX	/label= SCR3
XX	/note= "Structurally conserved region which is
XX	highly homologous to Ras-superfamily GEFs"
XX	
XX	WO200024768-A2.
XX	
XX	04-MAY-2000.

22-OCT-1999: 99WO-US24026.  
23-OCT-1998: 98US-0105507.  
16-NOV-1998: 98US-0108685.  
(MASI ) MASSACHUSETTS INST TECHNOLOGY.  
Kawasaki H, Graybiel A, Housman D;  
WPI: 2000-350690/30.  
N-PSDB: AAD00314.  
Isolated nucleic acid comprises nucleotide sequence encoding protein  
selected from normal or mutant CalDAG-guanine nucleotide exchange  
factor -  
Claim 64: Page 107-109; 128pp; English.  
The present sequence is a cyclic adenosine 3', 5',  
monophosphate-guanine nucleotide exchange factor1 (cAMP-GEF1) from rat.  
The cAMP-GEF1 has cAMP binding domain and Ras superfamily GEF  
domains. It has substrate specificity for Rap1a and is differentially  
distributed in brain and various other tissues. It selectively activates  
the Ras superfamily small G protein and Rap1a, functions as a negative  
regulator of Ras and directly couples the cAMP signal transduction  
system to Ras superfamily cascades. The present sequence  
is used for diagnosis and treatment (protein therapy) of  
cAMP-GEF-associated disorders. It is also useful for developing drugs  
and producing cell-lines or transgenic animal models for Ras-associated  
cancers.  
Sequence 884 AA;

Query Match	94.6%	Score 1664	DB 21	Length 884
Best Local Similarity	93.5%	Pred. No. 3.1e-159		
Matches 316	Conservative 11	Mismatches 11	Indels 0	Gaps 0
QY	1	MAALAEDEGWTKGQVLVYKNSAGDAIGIQDPARGVATSLGINERLFVYVNPQEVHELIPHP	60	
DB	547	MAALAHEDHMTKGOVLVYKNSAGDVYGIQDPARGVATSLGINERLFVYVNPQEVHELITHP	606	
QY	61	DQLPTVGSAGCLDLYSAKDLAQLTDHWSLFNSIHQVELIHVYLGPOHLRDYTTANLE	120	
DB	607	EQLPTLGSSMDLDYSAKDLAQLQTEHDHMLFNRIHQVELIHVYLGPOHLRDYTTANLE	666	
QY	121	RFMRRENFELQYVATVETCLCPVPCGPAOLLRKFTIKLAHLKQKLNSEFFAVMFCLSNSA	180	
DB	667	RFMRRENFELQYVATVETCLCPVPCGPAOLLRKFTIKLAHLKQKLNSEFFAVMFCLSNSA	726	
QY	181	ISRLAHTHERLPHVKRKILYSALERLDDSSMNHRYRLAKLSPVITFPMPLLLKDMFTI	240	
DB	727	ISRLAHTHERLPHVKRKILYSALERLDDSSMNHRYRLAKLSPVITFPMPLLLKDMFTI	786	
QY	241	HEGNHTIVENTINFEKPMRMAARAAMLHCRSHNVPILSPISRVSYSHLHEDSOVARISTC	300	
DB	787	HEGNHTIVENTINFEKPMRMAARAAMLHCRSHNVPILSPISRVSYSHLHEDSOVARISTC	846	
QY	301	SEOSTLSTRSPASTWAVYVQOLKVIDNQRELSTRSRELEP	338	
DB	847	SEOSTLSTRSPASTWAVYVQOLKVIDNQRELSTRSRELEP	884	
RESULT 4				
AAAY70968				
ID	AAAY70968	standard; Protein; 849 AA.		
XX	AAAY70968;			
XX				
XX	09-AUG-2000	(first entry)		
XX				
DE	Human Ras signalling pathway associated protein CAMP-GEFII.			
XX				
KW	Ras signalling pathway; CAMP-GEFII; cyclic adenosine monophosphate;			

[illegible]

Qy	121	REMRERFEDLYGWATATELCLCPVGPRAOILRKPFITAAIKOKKULNSFFAVMGSLNSA	180
Db	653	LFLRFRNEIDFWVWTEELICLCSQSKRVOLKFKFIIAACHKEYKNLNSFFAVMGSLNSIA	712
Qy	181	ISRLAHFWERIPKRVKRLYSALERLLDPSPMNHRYVRLALAKLSPVPIFMPDLLKDMTFI	240
Db	713	VSRLALTWELKPSKFKFYAEFFSLMDPSRNHRARLVYAKLEPPLIFPMULLIKDMTFI	772
Qy	241	HEGHNLTLEVLINFEKRMARARALHHCRRSHNPVPLSRVSLHEDSQVARISTC	300
Db	773	HEGKRTFLDNLVNFEEKRMILANTAFRVYRSQ--FFNPDOAAQNKKHHQDVR-----	822
Qy	301	SEQSLSTRSPASTMAYVOOLKLVNDNRSLRSLRELEP	338
Db	823	-----SYVRQLNVLIDNRQLTSQMSHRLRP	846
RESULT 5			
AAAY70967	ID	AAAY70967 standard; Protein; 291 AA.	
AAAY70967;	AC		
XX	XX		
XX	XX	09-AUG-2000 (first entry)	
DE	XX		
DE	Rat	Ras signalling pathway associated protein CAMP-GEFII.	
XX	XX		
KM	Rat;	Ras signalling pathway; CAMP-GEFII; cyclic adenosine monophosphate;	
KM	GEF;	guanine nucleotide exchange factor; Rap1A; diagnosis; treatment;	
KM	CAMP-GEF-	associated disorder; drug; transgenic animal model;	
KM	Ras-associated	cancer; protein therapy.	
XX			
OS	Rattus	novegicus.	
XX			
FH	Key	Location/Qualifiers	
FT	Region	48..71	
FT		/label= SCRI	
FT		/note="Structurally conserved region which is	
FT		highly homologous to Ras-superfamily GEFs"	
FT	Region	120..162	
FT		/label= SCR2	
FT		/note="Structurally conserved region which is	
FT		highly homologous to Ras-superfamily GEFs"	
FT	Region	199..220	
FT		/label= SCR3	
FT		/note="Structurally conserved region which is	
FT		highly homologous to Ras-superfamily GEFs"	
XX			
PN	WO200024768-A2.		
XX			
PD	04-MAY-2000.		
XX			
PE	22-OCT-1999;	99WO-US24826.	
XX			
XX	23-OCT-1998;	98US-0105507;	
PR	16-NOV-1998;	98US-0108685.	
XX			
PA	(MASI )	MASSACHUSETTS INST TECHNOLOGY.	
PI	Kawasaki H,	Graybiel A, Housman D;	
XX			
DR	WPI; 2000-350690/30.		
XX			
DR	N-PSDB; AAD00317.		
XX			
PT	Isolated	nucleic acid comprises nucleotide sequence encoding protein	
PT	selected	from normal or mutant Ca1DAG-guanine nucleotide exchange	
PT	factor	-	
XX			
XX			
PS	Claim 64;	Page 122; 128pp; English.	
XX			
CC	The	present sequence is a rat cyclic adenosine 3', 5'	
CC	monophosphate-guanine nucleotide exchange factorII (CAMP-GEFII)		
CC	which	has CAMP binding domain and Ras superfamily GEF	



PR 20-JUN-2000; 2000US-0598075.  
PR 19-JUL-2000; 2000US-0620325.  
PR 01-SEP-2000; 2000US-0654936.  
PR 15-SEP-2000; 2000US-0663561.  
PR 20-OCT-2000; 2000US-0693325.  
PR 30-NOV-2000; 2000US-0728422.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
PI Zhao QA, Wang D, Wehrman T, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wehrman T, Goodrich R,  
XX  
XX WPI; 2001-476283/51.  
DR N-PSDB; AAK51839.  
XX  
XX Nucleic acids encoding polypeptides with cytokine-like activities,  
PT useful in diagnosis and gene therapy -  
PS Claim 20; Page 3619-3620; 6221pp; English.  
XX  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAW8323-AAW80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activity/inhibit activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
CC (AAW0020) are omitted as the relevant pages from the sequence listing  
CC were missing at the time of publication.  
XX  
XX Sequence 456 AA:  
SQ  
Query Match 36.2%; Score 637.5; DB 22; Length 456;  
Best Local Similarity 42.0%; Pred. No. 1.3e-55;  
Matches 136; Conservative 60; Mismatches 99; Indels 29; Gaps 6;  
QY 15 VLKVNAGDAIGLQPDAGVATSLGLNRLFPVNPQEVNHELPHRDQGPVGSABGLD 74  
DB 157 ILVAVSSGSEKVLQPTEDCVFTALGINSHLFACSTRDSYEAIVLPPEETQVSPGDE-TH 215  
QY 75 LVSADKLAGQLTDHMSLFNSIHQVELIHVYL-GPQHLRDVTTANLFRMRFRNELQYVW 133  
DB 216 RVEPDEVANHLTAFFHMFRCVHELEFPDYVHGERGR--TANLELLQRCSEVTHW 273  
QY 134 ATELCLCPVPPRAQDLRRFKILAAHLKQKMLNSEFAVMFGLSNSAISRLAHWTWELPH 193  
DB 274 ATEVLLCEAPRGRAQDLKFKIKIAALCKQNDLSFEYAVMGIDNAVASRLRLTWELKPG 333  
QY 194 KYRKLYSALERLLDPSNMHRRVRLAKLSPPVIPPMLLDMTFIHGNTLVENLIN 253  
DB 334 KRNKLFKEFENTLDPGRNHSKYREVISKMKPPVIPPVPLLDLFTLHGSKTVLDGLVN 393  
QY 254 FEKRMAMARAARMLHRCSHNPVPLSLRSRVSHLHEDSQVARISPCSQSLSTRSPAST 313  
DB 394 IERKHSVAKKVTIRIKYRS-RPLCLD-MEASPNHL-----QT 428  
QY 314 WATVOQLKVIDNQRELISRLSRELE 337  
DB 429 KAVVROFQVIDNQNLLELSTYKLE 452

DT 22-OCT-2001 (first entry)  
XX  
XX Human polypeptide SEQ ID NO 2206.  
DE  
XX  
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200153312-A1.  
PN  
XX  
XX 26-JUL-2001.  
PD  
XX  
XX 26-DEC-2000; 2000WO-US34263.  
PF  
XX  
XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI; 2001-442253/47.  
DR N-PSDB; AAI58217.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
PT  
XX  
XX Example 4; SEQ ID NO 2206; 10078pp; English.  
PS  
XX  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAW38642-AAW42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC activation/inhibit activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
XX Sequence 456 AA:  
SQ  
Query Match 36.2%; Score 637.5; DB 22; Length 456;  
Best Local Similarity 42.0%; Pred. No. 1.3e-55;  
Matches 136; Conservative 60; Mismatches 99; Indels 29; Gaps 6;  
QY 15 VLKVNAGDAIGLQPDAGVATSLGLNRLFPVNPQEVNHELPHRDQGPVGSABGLD 74  
DB 157 ILVAVSSGSEKVLQPTEDCVFTALGINSHLFACSTRDSYEAIVLPPEETQVSPGDE-TH 215  
QY 75 LVSADKLAGQLTDHMSLFNSIHQVELIHVYL-GPQHLRDVTTANLFRMRFRNELQYVW 133  
DB 216 RVEPDEVANHLTAFFHMFRCVHELEFPDYVHGERGR--TANLELLQRCSEVTHW 273

Oy		134	A TELCLCPCVPGRRADLLKKEFTILANLHEQKNLSNFAMFGSLNSAISRATHTWERLP	193
Dd		274	A TEVELLCRAPGRCRADLLKKFKETIALCLQNODLLSEFYAAVMGDNAAVSRRLRTWEKLPD	333
Oy		194	KVRKYSALERLLDSNMHNRVRLAKLSPVIPMPPLLDMDPIHGNNHLVENLIN	253
Dd		334	KFNLFKRFENULTDCRNHKRSIREYISKMKPPVIYPVLPLDLDTFLHEGSKTLVGGLVN	393
Oy		254	EKKMBMARAAMLMHNCRSHNPVPLPSLRVSHLHEDSQVARISCTSEOSTRSPAST	313
Dd		394	IEKLHSVAEKVTIRKRS-RELCID-MEASPNHL-----QT 428	
Oy		314	WAYVOQLKYINDQRELRLSRELE	337
Dd		429	KAYNROFQVIDNQNLFFELSYKLE	452
 RESULT 9 ID AAE22105 standard; Protein: 246 AA.				
AC		AAE22105;		
XX		25-JUL-2002	(first entry)	
DE		Human	48921 guanine-nucleotide exchange cell phorbol-ester consensus.	
XX		Human:	GTP releasing factor: 48921; gene therapy: cytostatic; virucide;	
KW		neuroprotective:	vulnerary; hepatotropic; cardiant; hypotensive; cancer:	
KW		Inflammation:	diabetes; cellular proliferative disorder; differentiative:	
KW		brain:	liver; blood vessel; haematopoietic neoplastic; cerebrovascular;	
KW		leukemia;	meningitis; prion disease; Parkinson's disease; Pick disease;	
KW		Alzheimer's disease;	Huntington's disease; infection; hepatic injury;	
KW		Wilson's disease;	haemochromatosis; Gaucher's disease; atherosclerosis;	
KW		tumour;	hypertension; Kaposi sarcoma; chromosome mapping; tissue typing;	
KX		transgenic animal.		
OS		Homo sapiens.		
PN		MO200218577-AZ.		
PD		07-MAR-2002.		
XX		27-AUG-2001;	2001WO-US26696.	
Pf		30-AUG-2000;	2000US-228760P.	
PR		(MILL-) MILLENNIUM PHARM INC.		
PA		Curtis RAD;		
PI		WPI: 2002-351708/38.		
DR		New isolated human GTP releasing factor-48921 polypeptide for treating		
XX		cancer, inflammation, diabetes and pathogenic invasion of host cells		
PT		Disclosure; Fig 4; 110pp; English.		
XX		The invention relates to an isolated human GTP releasing factor, termed		
CC		48921 protein. Human 48921 DNA and protein are useful as diagnostic and		
CC		therapeutic agents for preventing a disease or condition associated with		
CC		an aberrant or unwanted 48921 activity in a subject, including cancer,		
CC		inflammation, diabetes and pathogenic invasion of host cells. 48921		
CC		protein is also useful for treating various disorders, including cellular		
CC		proliferative and differentiative disorders (metastatic, malignant,		
CC		cancer, hematopoietic neoplastic disorders including leukemias), brain		
CC		(cerebrovascular disorders, meningitis, Prion disease, pick disease,		
CC		Alzheimer's disease, Parkinson's disease, Huntington's disease, inborn		
CC		errors of metabolism), liver (viral, bacterial or parasitic infections,		
CC		hepatic injury, Wilson's disease, haemochromatosis, Gaucher's disease),		
CC		blood vessel (atherosclerosis, tumours, hypertension, Kaposi sarcoma and		
CC		pathology of therapeutic interventions in vascular diseases). 48921 DNA		
CC		and protein are useful in screening assays, detection assays (forensic		

Query Match	23.8%	Score 419.5	DB 23	Length 246
Best Local Similarity	43.3%	Pred. No. 6.4e-34		
Matches	91: Conservative	40: Mismatches	58: Indels	21: Gaps
73	LDL--VSAKKDLAGOLTDDHMSLEFNSHQVELIHYV-----LGPQHLNDVTANLERFMR	124		
DB	8 LDILHLDPHEELAEQTLTLDLFELEFKIEPSECGYWSNREKKGKENL---SPNLEKFTQ	63		
OY	125 RFNELQYVATELCLCPVGPFAOLRRKFIKLAHLAEQKNLSFFAVFGLSNSAISRL	184		
DB	64 RFNNISYVAATEILSESEKPPQRAKIKIEKFIKAQCRLENNNSLMATYSSGLNSSIYRL	123		
OY	185 AHTWERLPHKVKLYSALERLLDPSMNHRYRLALAKS-----PVIYPPMLLLKDM	237		
DB	124 KKTWEKVKERKKLFEELSELMDPSNNYKNYRELKSCGFVNHQSPPCIPIPLGYLKL	183		
OY	238 TFIHEGNTIYV--NLINFEKRMMAAKRAR	265		
DB	184 TFIHEGNPDYLDNTNLINFEKRMMLAKIIR	213		
RESULT 10				
AAB07792	ID	AAB07792	standard; Protein; 1499 AA.	
XX	AC	AAB07792;		
XX	DT	07-NOV-2000	(first entry)	
XX	DE	A murine guanine nucleotide releasing factor 4.		
XX	KW	guanine nucleotide releasing factor 4; GRF-4; Ras activator;		
KW	Nedd4 ubiquitination; cell metabolism; cell proliferation; cancer;			
KW	cell differentiation; cell transformation; neuronal disorder.			
XX	OS	Mus sp.		
XX	Key	Location/Qualifiers		
PH	Domain	135..253		
FT	Domain	/note="cNMP-binding domain"		
FT	Domain	266..322		
FT	Domain	/note="REM domain"		
FT	Domain	386..470		
FT	Domain	/note="PDZ domain"		
FT	Domain	594..692		
FT	Domain	/note="RA domain"		
FT	Domain	712..899		
FT	Region	/note="CDC25 domain"		
FT	Region	1403..1406		
FT	Region	/note="PY motif"		
FT	Region	1425..1428		
FT	Region	/note="PY motif"		
XX	PN	WO20043510-A2.		
XX	PD	27-JUL-2000.		
XX	PR	20-JAN-2000; 2000WO-CAL00042.		
XX	PR	20-JAN-1999; 99CA-2259830.		

[illegible][illegible]



Db 63 WATTELLKOTTLKPKRAEVLKFIYAKKRELNFNNSIMAIYVSSALSSSPISRLKKTWE 122  
 Oy 190 RLPKHVKRLYSALERLDPG-WNHRVYRLAL-----AKSPVYIPMPILLKDMFIHE 242  
 Db 123 KLSKTKKLPFEELLELDPSERNFKNYREALKSCNKPNOCPVPLGYLKDLPFIDE 182  
 Oy 243 GNHTLVEN-----LINFEMKMMARAARMLHCRSH-NPYLSPLRKSVSHLHDSOVARI 297  
 Db 183 GNPDPLENGTKGLVNEFKRKRIKILREIQLOSACOPYLKKNRNDIQELLRAS----- 237  
 Oy 298 STCSEOSLSTRSPASTWAVYQOLKVI-DNORELSRLSRELEP 338  
 Db 238 -----RPLEVLPEEDELELSELRIEP 259

RESULT 12  
 AAU99911  
 ID AAU99911 standard; Protein: 261 AA.  
 XX  
 AC AAU99911;  
 XX  
 DT 07-OCT-2002 (first entry)  
 XX  
 DE Human guanine nucleotide dissociation stimulator consensus sequence #2.  
 XX  
 KW 47476: guanine nucleotide dissociation factor infection;  
 KW haemopoietic disorder; blood clotting disorder; cancer;  
 KW autoimmune disorder; leukaemia; immunological disorder;  
 KW cardiovascular disorder; neurological disorder; cellular proliferation;  
 KW red blood cell disorder; viral disease; neurological disorder.  
 XX  
 OS Synthetic.  
 PN WO200240656-A2.  
 XX  
 PD 23-MAY-2002.  
 XX  
 PF 14-NOV-2001; 2001WO-US45291.  
 XX  
 PR 14-NOV-2000; 2000US-248331P.  
 PR 14-NOV-2000; 2000US-248362P.  
 PR 14-NOV-2000; 2000US-248365P.  
 PR 30-NOV-2000; 2000US-250077P.  
 PR 30-NOV-2000; 2000US-250176P.  
 PR 30-NOV-2000; 2000US-250327P.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Meyers RE, Curtis RAJ, Gluckmann MA;  
 XX  
 DR WPI; 2002-508325/54.  
 XX  
 XX Isolated 47476, 67210, 49875, 46842, 33201, 83378, 84233, 64708, 85041  
 PT or 84234 polypeptides, useful as reagents or targets for treating or  
 PT diagnosing pain or metabolic, liver, kidney, or cardiovascular  
 PT disorders  
 XX  
 PS Disclosure; Fig 2; 298pp; English.

This invention relates to the DNA and protein sequences of novel  
 isolated 47476, 67210, 49875, 46842, 33201, 83378, 84233, 64708, 85041  
 or 84234 proteins. The method of the invention is useful for treating a  
 disorder characterised by aberrant activity of 47476, 67210, 49875,  
 46842, 33201, 83378, 84233, 64708, 85041 or 84234-expressing cell in a  
 subject. The protein molecules can act as novel diagnostic targets and  
 therapeutic agents for controlling aberrant or deficient signal  
 transduction resulting, in e.g., haemopoietic disorders, including  
 blood clotting disorders, autoimmune disorders, or disorders related to  
 an inability to clear infections (e.g., viral or bacterial infections),  
 as well as disorders related to abnormal cellular proliferation or  
 differentiation, e.g., leukaemia. They may also be used to control  
 disorders of metabolic imbalance (e.g., disorders of lipopolysaccharide

CC biosynthesis or glycogen synthesis) immunological disorders,  
 CC cardiovascular disorders, neurological disorders, or cellular  
 CC proliferation and/or differentiation disorders, e.g., cancer, cell  
 CC motility and adhesion disorders, differential disorders, red blood  
 CC cell disorders, viral diseases, neurological disorders (e.g., brain  
 CC disorders), pain or metabolic disorders, liver disorders, kidney  
 CC imbalance, protein trafficking disorders and disorders of metal ion  
 CC bone metabolism. The sequences of the invention are also useful for  
 CC screening assays, predictive medicine (e.g., diagnostic assays,  
 CC prognostic assays, monitoring clinical trials, and pharmacogenetics);  
 CC and methods of treatment (e.g., therapeutic and prophylactic). The  
 CC present sequence represents a predicted consensus sequence motif  
 CC found in the protein of the invention.  
 CC  
 SO Sequence 261 AA;

Query Match 23.18; Score 405.5; DB 23; Length 261;  
 Best Local Similarity 37.6%; Pred. No. 1.e-32;  
 Matches 106; Conservative 41; Mismatches 92; Indels 43; Gaps 8;

Oy 75 LVSAKDLAQLTDHDMLEFNSIHVELIHVLPQHIRD---VTTALERFMRRNPLOY 131  
 Db 3 LDPEELAEQLTLDLDFELFRKIEPSELGSGWKRKSKSPSLAPQNLAEIFRNEVSN 62  
 Oy 132 WATTELLCPV--GPPAQLRKFKILAHLEKOKNLNFPFAVMGCLNSAISRLAHTWE 189  
 Db 63 WATTELLKOTTLKPKRAEVLKFIYAKKRELNFNNSIMAIYVSSALSSSPISRLKKTWE 122  
 Oy 190 RLPKHVKRLYSALERLDPG-WNHRVYRLAL-----AKSPVYIPMPILLKDMFIHE 242  
 Db 123 KLSKTKKLPFEELLELDPSERNFKNYREALKSCNKPNOCPVPLGYLKDLPFIDE 182  
 Oy 243 GNHTLVEN-----LINFEMKMMARAARMLHCRSH-NPYLSPLRKSVSHLHDSOVARI 297  
 Db 183 GNPDPLENGTKGLVNEFKRKRIKILREIQLOSACOPYLKKNRNDIQELLRAS----- 237  
 Oy 298 STCSEOSLSTRSPASTWAVYQOLKVI-DNORELSRLSRELEP 338  
 Db 238 -----RPLEVLPEEDELELSELRIEP 259

RESULT 13  
 ABB64459  
 ID ABB64459 standard; Protein: 1551 AA.  
 XX  
 AC ABB64459;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 20169.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 XX  
 DR N-PSDB; ABL08562.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -  
PS Disclosure: SEQ ID NO 20169; 21bp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB10176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins  
CC (AB101737-AB17072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1551 AA;  
  
Query Match 22.9%; Score 402.5; DB 22; Length 1551;  
Best Local Similarity 33.8%; Pred. No. 4.7e-31;  
Matches 110; Conservative 53; Mismatches 111; Indels 51; Gaps 10;  
  
QY 61 DOLGPVGSAGLDLV-----SAKDLAQLTDDHMSLFSNTHOVELIHVYLGPOH 110  
DB 842 DSTEPLVPDELALTELVRESNVHFLHNAVETALIOETLDP--FAVFRQIESTEVEDELFE 898  
QY 111 LBD-VTTANLEFRPMRFRNFOYVATELCICVPPRQALRKFKIKLAHLKEOKNLNSF 169  
DB 899 LMSRGVPMLSFAELVNMKFMVSVSEICAEHNIVRMKIVKOFIKIAHCKECPNFNSM 958  
QY 170 FAVMEGLNSAISRLAHTWERLPHVVKRLYSALERLDPSPMHHRYR-DATAKL--SPPV 226  
DB 959 FAIVSGLGGAVSRLRQWTEKLPSKYQRLFNDQLDMPSRMSKYRQIVSAELLAQHPI 1018  
QY 227 IFPMPLLDKMTFIEHGNTLVENLINFKMRMMAAAMLHCHSHNPVL-----278  
DB 1019 IFPPYIKKDLTFIHLGNDYRVDGLINFELKMLKEVLLTHMCS-SPYDLSTLELKG 1077  
QY 279 -SPLR-----SHVSHLEDQVARI-----STCSEOSTLRSP-----AST 313  
DB 1078 QSPSNALFSLNOMASQSNAAAGTYIAANAGQITTKRRKSTAAHPKMFEEAQMVRV 1137  
QY 314 MAYVOQLKVIDNQRELSRLSRELEP 338  
DB 1138 KAYLNSIKITLSDDELHKFLECEP 1162  
  
RESULT 14  
ABG66725  
ID ABG66725 standard; Protein; 1651 AA.  
XX  
AC ABG66725;  
XX  
DT 30-AUG-2002 (first entry)  
XX  
XX Human novel polypeptide #60.  
XX  
XX Human: inflammatory condition; shock; sepsis; immune response;  
KW cancer; wound healing; central nervous system disease; haematopoiesis;  
KW peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;  
KW myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;  
KW cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;  
KW bone degenerative disorder; periodontal disease; reperfusion injury;  
KW lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;  
KW allergic condition; thrombolysis; thrombosis; coagulation disorder;  
KW fungal infection.  
XX  
XX Homo sapiens.  
XX  
XX WO200244340-A2.

PD 06-JUN-2002.  
XX  
XX 30-NOV-2001; 2001WO-US47004.  
XX  
XX 30-NOV-2000; 2000US-0028952.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YF, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;  
PI Yamazaki V, Ujwal ML, Drmanac RT;  
XX  
XX WPI: 2002-508509/54.  
DR N-PSDB: ABR94949.  
XX  
XX Novel nucleic acids and polypeptides for diagnosis, treatment of  
PT inflammatory, autoimmune, nervous system, myeloid or lymphoid cell  
PT disorders, cancer and promoting wound healing -  
XX  
XX Claim 10; Page 629-632; 672pp; English.  
XX  
XX The invention relates to human novel polynucleotides and associated  
CC polypeptides. The polynucleotides and polypeptides are useful for  
CC treating inflammatory conditions such as arthritis, nephritis, Crohn's  
CC disease, ischaemia-reperfusion injury, shock, sepsis, immune responses  
CC and cancer and for promoting wound healing. The sequences are used to  
CC induce the proliferation of neural cells and regeneration of nerve and  
CC brain tissue, and are useful for the treatment of central and peripheral  
CC nervous system diseases and neuropathies, such as Alzheimer's disease,  
CC Parkinson's disease, Huntington's disease and amyotrophic lateral  
CC sclerosis. The sequences are involved in chemoattract or chemokinetic  
CC activity, regulation of haematopoiesis, treatment of myeloid or lymphoid  
CC cell disorders and platelet disorders such as thrombocytopenia,  
CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue  
CC growth, tissue repair, healing of burns, incisions, ulcers, treatment of  
CC osteoporosis, osteoarthritis, bone degenerative disorders and periodontal  
CC disease. The sequences of the invention are also useful for gut  
CC protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues, immune deficiencies and disorders  
CC including severe combined immunodeficiency (SCID), bacterial or fungal  
CC infections, autoimmune disorders e.g. multiple sclerosis and myasthenia  
CC gravis, allergic conditions such as asthma, thrombolytic or thrombosis  
CC and coagulation disorders. Sequences ABG66666-ABG66756 represent human  
CC novel polypeptides of the invention.  
XX  
SQ Sequence 1651 AA;  
  
Query Match 22.2%; Score 390.5; DB 23; Length 1651;  
Best Local Similarity 30.0%; Pred. No. 8.5e-30;  
Matches 101; Conservative 57; Mismatches 128; Indels 51; Gaps 5;  
  
QY 35 VATSLGLNERLEFVNPQEVHLLIPHPDQGLPTVGSAGELDLVSAKDLAQLTDDHMSLFN 94  
DB 868 LADRIQLNGRYVILKNNMETFLCSDEDOELVYKESQLMLQSTLEVATOLSMRFDLFR 927  
QY 95 SIHOVELIHVYLGPOHLDVY-TANLEFRMFRNFOYVATELCICVPPRQALRKFE 153  
DB 928 NIEPTEYIDDLF--KLMSKTGNTHLKRFEEDIVDETGVASSELITTEANOLKRIKHF 984  
QY 154 IKLAHLKEOKNLNSFVAFMGLSNSAISRLAHTWERLPHVVKRLYSALERLDPSPMHHR 213  
DB 965 IKIALHCKECPNFNSMFAIISGLNLAIVARLGRTEKLPSTREKHLDQLDIPDRSRMA 1044  
QY 214 VYR--LALAKSPPVIPPMPLLDKMTFIEHGNTLVENLINFKMRMMAAAMLHCHCR 271  
DB 1045 KYRNITLSSQSQMPRIPLFPVYKDMFTLHGNDKSKVDGLNVEFLRMISREIROVYVMT 1104  
QY 272 SHNPVPLSPPLSR-----VSHLEDQVARI-----301  
DB 1105 SANMDPAMMFQROSLSSQSTNSNMLDVGGAHKRRARRSSLLNAKKIYEDQAMAR-----1159  
QY 302 EGSLSRSPASTWAVVOQLKVIDNQRELSRLSRELEP 338  
DB 1160 -----KVQYLSLSDVETDEKFOMSLOWEP 1186

RESULT 15  
 AAB42658  
 ID AAB42658 standard; Protein: 1675 AA.  
 XX  
 AC AAB42658;  
 XX  
 DT 08-FEB-2001 (first entry)  
 DE Human ORFX ORF2422 polypeptide sequence SEQ ID NO:4844.  
 XX  
 KW Human; Open reading frame; ORFX; detection; cytosolic; hepatotropic;  
 KW vlnarary; antiparaksinon; antiparaksinon; noctropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antilethritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antilethritic;  
 KW antilethritic; antilethritic; antilethritic; antilethritic;  
 KW antilethritic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.  
 XX  
 OS Homo sapiens.  
 PN WO200058473-A2.  
 PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000MO-US08621.  
 XX  
 PR 31-MAR-1999; 99US-0127607.  
 PR 02-APR-1999; 99US-0127636.  
 PR 05-APR-1999; 99US-0127728.  
 PR 30-MAR-2000; 2000US-0540763.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shinkets RA, Leach M;  
 PI  
 DR N-PSDB; AAC76867.  
 DR  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -  
 XX  
 PS Claim 11: Page 4028-4032; 5507pp; English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytosolic; hepatotropic; vlnarary;  
 CC antiparaksinon; antiparaksinon; noctropic; neuroprotective;  
 CC osteopathic; anticonvulsant; antilethritic; immunosuppressant;  
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
 CC antilethritic; antilethritic; antilethritic; antilethritic;  
 CC antilethritic; antilethritic; antilethritic; antilethritic;  
 CC the presence of or predisposition to, or preventing or treating  
 CC pathological conditions associated with an ORFX-associated disorder. The  
 CC nucleic acids can be used to express ORFX proteins in gene therapy  
 CC vectors. The proteins and nucleic acids may be used to treat cancers,  
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance

CC coagulation; to inhibit thrombosis; and as a contraceptive.  
 XX  
 SQ Sequence 1675 AA;  
 Query Match 22.28; Score 390.5; DB 21; Length 1675;  
 Best Local Similarity 30.04; Pred. No. 8.7e-30;  
 Matches 101; Conservative 57; Mismatches 128; Indels 51; Gaps 5;  
 QY 35 VATSIGNERLFVNNQEVNELIRHPDQLPTVGSAGLDLVSAKDLDTDHMSLFN 94  
 DB 891 LADRIDNGRYVYKNNMETETLCSDEDAELVESQLSMQLSTIEVATQLSMKDFLFR 950  
 QY 95 SIHQVELIHYVLGPOLRDVT-TANLERFMRRENELOVYVATECLCPVGPRAQLRKF 153  
 DB 951 NIEPTVIDDLF---KLNSTGNTHLKRFEDIYNGEFTFWASLITFANOLKMKIKHF 1007  
 QY 154 IKLAHLKEOKNLNSFPVAVFGLSNSAISRLAHTWERLPHKVKRLYSALERLLDPSVNR 213  
 DB 1008 IKTALHCRECKNFSMFAISGLNLASVARTRGTEKLPSEKYEKHLQDIDPDSRNMA 1067  
 QY 214 VYR--TALAKLSPVITPEMPLLLKDMFTIEGNTLVENLINFEMRMARARMLHCR 271  
 DB 1068 KYNNILSSQSMQPIILPFVVKMDTFLEHGDSKVDGLVNEKLMISKEIRQVVRMT 1127  
 QY 272 SHNPVPLSPLSR-----VSHLHEDSQVARISTCS 301  
 DB 1128 SANMDPAMFROSLSGSTNSMMDVQGAHKKRARSSLLNAKKRIEDAGMAR----- 1182  
 QY 302 EOSLSTRSPASTWAVYQOLKVIDNORELSRLSRELP 338  
 DB 1183 -----KVKQYLSLDVETDEKFKQMSIQWER 1209

Search completed: October 7, 2003, 07:13:10  
 Job Time : 88 secs

**THIS PAGE BLANK (USPTO)**